

# Semiautomatische Detektion von Skelettbefall im PET/CT bei Kindern und Jugendlichen mit Hodgkin-Lymphom

Dissertation  
zur Erlangung des akademischen Grades  
Dr. med.

an der Medizinischen Fakultät  
der Universität Leipzig

eingereicht von:  
Maximilian Epstude

Geburtsdatum / Geburtsort:  
27.01.1988, Jena

angefertigt an:  
Klinik und Poliklinik für Nuklearmedizin  
Universitätsklinikum Leipzig

Betreuer:  
Frau Professor Dr. med. habil. Regine Kluge  
Herr Dr. med. Lars Kurch

Beschluss über die Verleihung des Doktorgrades vom:  
21.05.2019

# Inhaltsverzeichnis

1	Abkürzungsverzeichnis .....	1
2	Bibliographische Angabe .....	4
3	Einleitung .....	5
	3.1 FDG-PET und deren besondere Bedeutung in der Onkologie .....	5
	3.2 Verschiedene Methoden der FDG-PET-Auswertung .....	6
	3.3 Therapieoptimierungsstudien zur Behandlung des Hodgkin-Lymphoms im Kindes- und Jugendalter .....	8
	3.4 Ermittlung von Risikofaktoren zur Therapietitration und besondere Bedeutung der FDG-PET/CT .....	10
	3.5 Organisationsstruktur der EuroNet-PHL-C1-Studie und Einsatz der PET/CT .....	11
	3.6 Detektion von Skelettbefall .....	12
	3.7 Bedeutung von Computerprogrammen für die bildgebende Diagnostik, insbesondere für die Detektion eines Skelettbefalls.....	14
	3.8 Nicht-computergestützte, visuelle Bewertung des Skelettsystems im Rahmen der EuroNet-PHL-C1-Studie .....	18
	3.9 Bisherige Erkenntnisse zur Diagnostik von Skelettbefall mittels PET/CT beim pädiatrischen HL .....	20
4	Aufgabenstellung .....	23
5	Materialien und Methoden .....	24
	5.1 Stand des Computerprogrammes zu Beginn der Doktorarbeit und Optimierung seiner Funktionsweise .....	24
	5.1.1 Benutzeroberfläche .....	24
	5.1.2 Extraktion des Skelettsystems aus dem PET-Datensatz über eine Dichteschwelle der CT-Komponente und methodenimmanente Probleme....	25
	5.2 Experimentelle Ansätze zur softwaregestützten Detektion tumorbefallener Skelettanteile in der PET nach Programmoptimierung .....	29
	5.2.1 Ansatz I: Analyse des Nuklidspeicherverhaltens im virtuell extrahierten Skelett durch Normierung auf das Speicherverhalten im Referenzorgan Leber .....	29
	5.2.2 Ansatz II: Histogrammanalyse.....	32

5.2.3 Ansatz III: Aktivitätsanalyse ausgewählter Skelettregionen mit und ohne Lymphommanifestationen.....	37
5.2.4 Ansatz IV: Vergleich der Aktivitätsbelegung von Wirbelkörpern und Wirbelkörperläsionen .....	39
5.2.5 Ansatz V: Test verschiedener Standardabweichungen über dem Mittelwert der Aktivitätsbelegung des Gesamtskeletts bzw. eines physiologischen Wirbelkörperuptakes als Grundlage für die programmgestützte Läsionendetektion.....	42
5.2.6 Überprüfung des als am besten geeigneten Algorithmus an Patienten mit physiologischem Skelettuptake.....	46
5.2.7 Überprüfung und Optimierung des Größenkriteriums zur Anwendung auf Datensätze mit unterschiedlicher Matrixgröße .....	46
5.3 Evaluation des experimentell gefundenen, besten Programmalgorithmus an einer großen Patientenpopulation zur Detektion und Quantifizierung des Skelettbefalls mittels PET .....	47
5.4 Korrelation des durch den Algorithmus quantifizierten Skelettbefalls mit laborchemischen und klinischen Outcomedaten der Patienten .....	53
6 Ergebnisse .....	54
6.1 Optimierung der Funktionsweise der vorbestehenden Programmversion (Extraktion des Skelettsystems aus dem PET-Datensatz aus der CT-Komponente) .....	54
6.2 Experimentelle Ansätze zur softwaregestützten Detektion von Skelettbefall in der PET nach Programmoptimierung.....	56
6.2.1 Ansatz I: Analyse des Nuklidspeicherverhaltens im extrahierten Skelett unter Normierung auf das Referenzorgan Leber.....	56
6.2.2 Ansatz II: Histogrammanalyse .....	61
6.2.3 Ansatz III: Aktivitätsanalyse ausgewählter Skelettregionen mit und ohne Lymphommanifestationen.....	65
6.2.4 Ansatz IV: Vergleich der Aktivitätsbelegung von Wirbelkörpern und Wirbelkörperläsionen .....	74
6.2.5 Ansatz V: Test verschiedener Standardabweichungen über dem Mittelwert der Aktivitätsbelegung des Gesamtskeletts bzw. des physiologischen	

Wirbelkörperuptakes als Grundlage für die programmgestützte Läsionendetektion.....	77
6.2.6 Überprüfung des als optimal getesteten Algorithmus an Patienten mit physiologischem Skelettuptake.....	80
6.2.7 Überprüfung und Optimierung des Größenkriteriums zur Anwendung auf Datensätze mit unterschiedlicher Matrixgröße .....	82
6.3 Evaluation des experimentell ermittelten besten Programmalgorithmus zur Quantifizierung des Skelettbefalls in der FDG-PET an einer großen Patientenpopulation .....	84
6.4 Korrelation des durch den Algorithmus quantifizierten Skelettbefalls mit laborchemischen und klinischen Outcomedaten der Patienten .....	97
7 Diskussion .....	100
7.1 Extraktion des Skelettsystems aus dem PET und dabei auftretende methodenimmanente Probleme .....	100
7.2 Experimentelle Ansätze zur Detektion von Skelettbefall in der PET nach erfolgter Programmoptimierung .....	103
7.2.1 Ansatz I: Analyse des Nuklidspeicherverhaltens im extrahierten Skelett unter Normierung auf das Referenzorgan Leber.....	103
7.2.2 Ansatz II: Histogrammanalyse.....	105
7.2.3 Ansatz III: Aktivitätsanalyse ausgewählter Skelettregionen mit und ohne Lymphommanifestationen.....	108
7.2.4 Ansatz IV: Charakterisierung und Vergleich der Aktivitätsbelegungen physiologischer Wirbelkörper und Wirbelkörperläsionen .....	109
7.2.5 Test verschiedener Standardabweichungen über dem Mittelwert der Aktivitätsbelegung des Gesamtskeletts bzw. eines physiologischen Wirbelkörpers als Grundlage für die Läsionendetektion.....	111
7.3 Evaluation des experimentell gefundenen besten Programmalgorithmus zur Quantifizierung des Knochenmarkbefalls an einer großen Patientenpopulation .....	115
7.3.1 Sensitivität des Algorithmus .....	115
7.3.2 Verteilung der Skelettläsionen und Umfang der CT-Skelettmaske.....	118
7.3.3 Einschränkungen und Verbesserungsmöglichkeiten.....	119

7.4	Bedeutung (semi-)automatischer Algorithmen in der funktionellen Bildgebung und Relevanz von hämatogenem Skelettbefall beim Hodgkin-Lymphom und soliden Tumoren .....	123
7.5	Weitere Anwendungsmöglichkeiten und mögliche Bedeutung des entwickelten Algorithmus in der klinischen Routine .....	132
8	Zusammenfassung der Arbeit.....	135
9	Literaturverzeichnis.....	140
10	Abbildungsverzeichnis .....	148
11	Tabellenverzeichnis .....	150
12	Erklärung über die eigenständige Abfassung der Arbeit.....	152
13	Lebenslauf .....	153
14	Danksagung.....	155

# **1 Abkürzungsverzeichnis**

2D	zweidimensional
3D	dreidimensional
Abb.	Abbildung
AUC-CSH	Area Under The Curve – Cumulative SUV histogram
BSI	Bone Scan Index
BWS	Brustwirbelsäule
bzw.	beziehungsweise
ca.	circa
CD	Compact Disc
<sup>11</sup> C-PIB-PET	Kohlenstoff-11-Pittsburgh-Compound-B-Positronenemissionstomografie
CT	Computertomografie
DGN	Deutsche Gesellschaft für Nuklearmedizin
d. h.	das heißt
Dr. med.	Doctor medicinae
Dr. med. habil.	Doctor medicinae habilitatus
DVD	Digital Versatile Disc
EANM	European Association of Nuclear Medicine
EARL	EANM Research Limited
E-Befall	Lymphomwachstum in das Skelett per continuitatem
EFS	Event Free Survival (Ereignisfreies Überleben)
engl.	englisch
ERA	Early Response Assessment (Frühresponsekontrolle)
ERA-PET	Frühresponse-Positronenemissionstomografie
et al.	et alii
EU	Europäische Union
evtl.	eventuell
<sup>18</sup> F-FDG	Fluor-18-Fluordesoxyglukose
FDG-PET	Fluordesoxyglukose-Positronenemissionstomografie
<sup>18</sup> F-FLT	Fluor-18-Fluorthymidin
FLAB	Fuzzy Locally Adaptive Bayesian
FN	falsch-negativ
g	Gramm
ggf.	gegebenenfalls
GIST	Gastrointestinaler Stromatumor
GLUT	Glukosetransporter
GPOH	Gesellschaft für Pädiatrische Onkologie und Hämatologie
Gpt	Gigapartikel
HD	Hodgkin's Disease
HL	Hodgkin-Lymphom
HU	Hounsfield Unit
HWS	Halswirbelsäule
ICRP	International Commission on Radiological Protection
ID	Identifikationsnummer
Ig	Immunglobulin
IgA	Immunglobulin A
IgG	Immunglobulin G
IgM	Immunglobulin M

I	Liter
log	Logarithmus
KM	Knochenmark; Kontrastmittel
LDH	Laktatdehydrogenase
<sup>177</sup> Lu-PSMA	Luthetium-177-Prostate-Specific-Membrane-Antigen
LWK	Lendenwirbelkörper
LWS	Lendenwirbelsäule
max.	maximal
MBP	Mediastinaler Blutpool
MIP	Maximum-Intensitäts-Projektion
ml	Milliliter
MRT	Magnetresonanztomografie
MTV	Metabolisches Tumorvolumen
OEPA	Oncovin Etoposid Prednison Adriamycin
o. g.	oben genannte
OS	Overall Survival
Pat.	Patient
Pat.-ID	Patientenidentifikationsnummer
PET	Positronenemissionstomografie
PET/CT	Positronenemissionstomografie/Computertomografie
PET/MRT	Positronenemissionstomografie/Magnetresonanztomografie
PHL	Pädiatrisches Hodgkin-Lymphom
Prof.	Professor
PVE	Partialvolumeneffekt
qPET-Wert	Quantitativer Positronenemissionstomografie-Wert (Quotient aus SUV <sub>peak</sub> einer (nodalen) Hodgkin-Läsion und SUV <sub>mean</sub> eines Leberquaders)
qPET <sub>max</sub> -Wert	Maximaler qPET-Wert
qPET <sub>mean</sub> -Wert	Mittlerer qPET-Wert
Rel.	Relative
ROI	Region of Interest
ROVER	Region of interest (ROI) visualization, evaluation and image registration
S.	Seite
SD	Standard Deviation (Standardabweichung)
SD <sub>SUV</sub>	Standardabweichung des SUV
SD <sub>WK</sub>	Standardabweichung des Wirbelkörpers
s. g.	so genannte
SMN	Sekundäre maligne Neoplasien
SPECT	Single-Photon-Emissions-Computertomografie
SST	Sekundäre solide Tumoren
Std.-Abw.	Standardabweichung
SUV	Standard Uptake Value
SUV <sub>avg</sub>	(durchschnittlicher) SUV average
SUV <sub>bw</sub>	(körpergewichtsadaptierter) SUV body weight
SUV <sub>max</sub>	Maximaler SUV
SUV <sub>mean</sub>	Durchschnittlicher SUV
SUV <sub>median</sub>	Medianer SUV
SUV <sub>peak</sub>	„Spitzen-SUV“ (Mittelwert des SUV der vier höchsten Voxel)
SUV <sub>search</sub>	SUV einer Suchregion

Tab.	Tabelle
TOS	Therapieoptimierungsstudie
v. a.	vor allem
vgl.	vergleiche
VOI	Volume of Interest
VRT	Volume Rendering Technique
vs.	versus
WK	Wirbelkörper
z. B.	zum Beispiel
z. T.	zum Teil



## **2 Bibliographische Angabe**

Im Rahmen dieser Arbeit wurde ein Algorithmus zur semiautomatischen Detektion von pathologischen, skelettalen Mehranreicherungen in der  $^{18}\text{F}$ -FDG-PET entwickelt, der gleichzeitig als Grundlage für die Ermittlung des metabolischen, skelettalen Tumorstadiums diene.

Hierzu wurden PET/CT-Datensätze von Kindern und Jugendlichen, die an einem Hodgkin-Lymphom erkrankt waren und im Zeitraum von 2007 bis 2013 in der EuroNet-PHL-C1-Studie behandelt wurden, analysiert. Bei 142 Kindern war im Rahmen einer zentralen Referenzbeurteilung im PET/CT ein Skelettbefall diagnostiziert und dokumentiert worden. Auf dieser Basis (Goldstandard) erfolgte die Testung verschiedenster Methoden, von denen sich die nachfolgende als die am besten geeignete erwies:

Als Referenzregion zur Bestimmung eines physiologischen Stoffwechselniveaus im Skelett ( $\text{SUV}_{\text{mean}}$ ) wurde eine 15 bis 30 ml große VOI in einem nach visueller Einschätzung nicht tumorbefallenen Wirbelkörper (meist LWK 4) platziert. Auf Grundlage des damit ermittelten physiologischen Skelettstoffwechselniveaus identifizierte das Suchprogramm skelettale Läsionen, sofern diese einen  $\text{SUV} > (\text{SUV}_{\text{mean}} + 2,5 \text{ SD})$  aufwiesen. Dieser in das Programm implementierte Algorithmus wurde an den PET/CT-Datensätzen der 142 Patienten mit visuell detektiertem Skelettbefall validiert.

Dabei wurde ein Skelettbefall bei 130 von 142 Patienten (Sensitivität auf Patientenebene von 91,5 Prozent) programmbasiert korrekt diagnostiziert. Von 1015 visuell erfassten skelettalen Läsionen wurden 774 durch das Suchprogramm richtig erkannt (Sensitivität auf Regionenebene von 76,3 Prozent). Von 5375 nicht befallenen Regionen wurden 5137 korrekterweise auch nicht als Läsionen durch das Programm angezeigt (Spezifität auf Regionenebene von 95,6 Prozent). Das Ausmaß des Skelettbefalls in Form des metabolischen, skelettalen Tumorstadiums wurde mit dem ereignisfreien Überleben und dem Gesamtüberleben verglichen: Es ergab sich keine Korrelation.

Die Arbeit umfasst 156 Seiten, 59 Abbildungen, 31 Tabellen und 96 Literaturangaben.

### **3 Einleitung**

#### **3.1 FDG-PET und deren besondere Bedeutung in der Onkologie**

Tumorzellen weisen einen erhöhten Glukosestoffwechsel auf (Warburg 1956), was u. a. auf eine vermehrte Expression von Glukosetransporter 1 (GLUT 1) auf der Zelloberfläche zurückgeht (Hoekstra et al. 2000), durch welchen Glukose in die Zelle transportiert wird. Die Eigenschaft des erhöhten Glukoseverbrauchs vitaler Tumorzellen macht man sich diagnostisch bei der Positronenemissionstomografie (PET) zunutze: An ein Glukosemolekül wird ein Positronenstrahler (in der Regel  $^{18}\text{F}$ ) gekoppelt, es entsteht  $^{18}\text{F}$ -Fluordesoxyglukose ( $^{18}\text{F}$ -FDG). Zur Erstellung eines Positronenemissionstomogramms werden dem Patienten geringe Mengen an  $^{18}\text{F}$ -FDG intravenös verabreicht. Dieses reichert sich sodann überproportional häufig in pathologisch veränderten Zellen mit gesteigertem Glukosestoffwechsel an (Ruhlmann et al. 1998).

Nach der Aufnahme in die Zelle durchläuft  $^{18}\text{F}$ -FDG zunächst den ersten Schritt der Glykolyse, wird also zum größten Teil phosphoryliert. Nach der Phosphorylierung können die weiteren Schritte der Glykolyse aufgrund der Fluoridierung des Glukosemoleküls nicht durchlaufen werden,  $^{18}\text{F}$ -FDG akkumuliert in der Zelle (Trapping). Die Akkumulation von  $^{18}\text{F}$ -FDG an Stellen im Körper mit erhöhtem Glukoseumsatz kann wiederum von außen durch die von  $^{18}\text{F}$  emittierten Positronen sichtbar gemacht werden. Positronen reagieren nämlich als positiv geladene Teilchen mit den negativ geladenen Elektronen des Umgebungsgewebes durch Vernichtung ihrer selbst (Annihilation). Hierbei entstehen zwei Gammaquanten, die im Winkel von  $180^\circ$  auseinanderfliegen und von Detektorsystemen registriert werden können. Die anschließende und technisch aufwendige Rekonstruktion der aufgenommenen Impulse führt zu einem für den Arzt auswertbaren Tomogramm.

Nach (Wahl et al. 2009) können solide Tumore erst klinisch diagnostiziert werden, wenn sie eine Masse von 10 bis 100 g erreichen (ca.  $10^{10}$  bis  $10^{11}$  Tumorzellen). Heutige PET-Systeme können Tumoren ab einer Größe von 0,4 cm erfassen, was einer Tumormasse von etwa 0,1 bis 0,5 g (ca.  $10^8$  Zellen) entspricht.

### **3.2 Verschiedene Methoden der FDG-PET-Auswertung**

PET-Bilder können qualitativ, semiquantitativ und quantitativ ausgewertet werden.

Bei der qualitativen Analyse wird der Kontrast zwischen einer Läsion und dem umgebenden Gewebe visuell verglichen. Verschiedene Autoren betonen den Nutzen der qualitativen Analyse für die Bewertung von Therapieansprechen und zur Rezidivdiagnostik (Poeppel et al. 2009). (Itti et al. 2010) zeigten z. B. bei Patienten mit diffusem großzelligen Lymphom, dass ein residueller Glukosestoffwechsel unterhalb des Niveaus der Leber mit einem um 30 Prozent verbesserten ereignisfreien Überleben einhergeht. Der rein visuelle Vergleich mit einer Referenzregion (z. B. mediastinaler Blutpool [MBP], Leber oder Hirn) führt aber nachweislich zu einer hohen Interobservervariabilität, was wiederum nur eine eingeschränkte Vergleichbarkeit der qualitativen Analyse erlaubt (Hoekstra et al. 2000). Für Lymphome ist die visuelle Bewertung des Reststoffwechsels mithilfe einer 5-Punkte-Skala (Deauville-Score, Abb. 1) seit 2009 internationaler Standard (Meignan et al. 2009). Hierbei dienen der mediastinale Blutpool und die Leber als Referenzbezirke.

1	Kein Uptake
2	Uptake < Mediastinum
3	Uptake > Mediastinum aber < Leber
4	Uptake moderat über dem der Leber
5	Uptake deutlich über dem der Leber

Abb. 1: Deauville-5-Punkte-Skala für PET-Scans (Meignan et al. 2009)

Aber auch die Anwendung des Deauville-Scores birgt Fehlerquellen, die wiederum Grund für eine erhöhte Interobservervariabilität sein können (Hasenclever et al. 2014). Bei Anwendung semiquantitativer Methoden kann im Gegensatz zur visuellen Bewertung die Impulsdichte in einem Herd in Zahlen ausgedrückt werden, was eine quantitative Beschreibung sowohl der Intensität als auch der Ausdehnung des pathologischen Befundes ermöglicht.

Die semiquantitative Analyse erfordert als Grundvoraussetzung ein dreidimensionales, schwächungskorrigiertes Bild. Ferner bedarf die semiquantitative Analyse regelmäßiger Kalibrierungen des Scannersystems. Diese erfolgen mithilfe von s. g. Phantommessungen. In dem Phantom werden definierte Aktivitätskonzentrationen platziert, welche korrekt durch das Scannersystem abgebildet werden müssen. Ist dies nicht der Fall, sind entsprechende Eichungen, selten auch Reparaturen, erforderlich, um die gewünschte Genauigkeit wiederzuerlangen.

Im Zusammenhang mit der semiquantitativen Analyse ist der Standard Uptake Value (SUV) der wichtigste Parameter. Dieser drückt die relative Gewebekonzentration des Tracers als Zahl aus. Dabei wird die aus den korrigierten PET-Bildern ermittelte Impulsdichte durch die applizierte Aktivitätsmenge dividiert und mit einem Normierungsfaktor multipliziert (Zasadny und Wahl 1993). Im klinischen Alltag und somit auch in vielen Studien hat sich der s. g. SUV<sub>bw</sub> (body weight) bewährt. Hierbei dient das Körpergewicht des Patienten als Normierungsfaktor. Insbesondere in Multizenterstudien sind SUV-Werte, die von unterschiedlichen PET-Einrichtungen mit hinsichtlich Kalibrierung nicht aufeinander abgestimmten Scannersystemen generiert wurden, nur bedingt vergleichbar. Das liegt daran, dass der SUV-Wert neben dem Körpergewicht auch von der Körpergröße, dem Blutglukosespiegel, Partialvolumeneffekten, der Art des Scanners und der durchgeführten Rekonstruktion beeinflusst wird (Keyes 1995). Auch der Aufnahmezeitpunkt nach Radiotracerapplikation und die damit verbundene Bildqualität können einen Einfluss auf den SUV-Wert haben (Brown et al. 2010). Um eine Vergleichbarkeit semiquantitativer Analysen innerhalb von Multizenterstudien zu gewährleisten, kann der PET-Scanner auf ein einheitliches Phantom geeicht werden. Dies geschieht z. B. im Rahmen der von der European Association of Nuclear Medicine (EANM) initiierten EARL-Zertifizierung, an der verschiedene Kliniken europaweit teilnehmen. Allerdings ist die Teilnahme nicht zwingend vorgeschrieben. Um dennoch multizentrisch vergleichbare, semiquantitative Analysen durchführen zu können, kann man sich eines Kunstgriffs bedienen, bei dem der SUV<sub>peak</sub> des Reststoffwechsels durch den SUV<sub>mean</sub> innerhalb eines standardisierten Volumens, dem Volume of Interest (VOI), in einem Referenzorgan dividiert wird (Hasenclever et al. 2014). Durch die Quotientenbildung werden weitgehend die geräteimmanenten Einflussfaktoren auf den SUV ausgeschaltet. Somit bleiben, wie auch bei EARL-zertifizierten Geräten, nur noch untersuchungsspezifische Einflussgrößen wie Dauer zwischen Tracerapplikation und Scanbeginn. Letztere sollten jedoch bei Einhaltung von Leitlinien und/oder Studienprotokollkriterien keinen überbordenden Effekt haben.

Die genaueste Approximation an die tatsächlichen Verhältnisse in vivo erhält man durch die dritte Möglichkeit der Datenanalyse: die quantitative Methode. Hierfür sind dynamische Aufnahmen über einen längeren Zeitraum mit mehreren arteriellen Blutentnahmen im zeitlichen Verlauf erforderlich. Aus diesen Parametern kann unter Zuhilfenahme von Kompartimentmodellen eine im Vergleich zur semiquantitativen

Methode exaktere Berechnung der Höhe des Glukosemetabolismus erfolgen. Eine quantitative Analyse ist allerdings im klinischen Alltag nicht praktikabel und wird deshalb vor allem in der Forschung angewandt (Castell und Cook 2008).

### **3.3 Therapieoptimierungsstudien zur Behandlung des Hodgkin-Lymphoms im Kindes- und Jugendalter**

Hodgkin-Lymphome (HL) im Kindes- und Jugendalter sind seltene Erkrankungen, die für etwa 6 Prozent aller pädiatrischen Malignome verantwortlich sind (Rathore et al. 2012). In der Altersgruppe der 15-Jährigen bis 19-Jährigen ist das Hodgkin-Lymphom jedoch die häufigste maligne Erkrankung (Simpson et al. 2008). Die Fünf-Jahres-Überlebensrate ist von etwa 80 Prozent im Jahre 1975 auf aktuell über 95 Prozent angestiegen (Smith et al. 2010). Durch kumulative Effekte von therapeutisch verabreichter Radiotherapie und Polychemotherapie werden allerdings verschiedene Spätschäden verursacht. So bedingen Agentien der Polychemotherapie wie Alkylantien die Entwicklung von Infertilität, Anthrazykline tragen ein kardiotoxisches Potenzial und Prednisolon kann u. a. zum Auftreten aseptischer Knochennekrosen führen (Dörffel et al. 2010). Besonders bedenklich ist aber die Induktion von Zweitmalignomen: Die eher durch die Polychemotherapie verursachten Leukämien treten selten und in der Regel kurze Zeit nach der Behandlung auf und erreichen nach 10 bis 15 Jahren ein Plateau in ihrer kumulierten Inzidenz (Daw et al. 2010). Im Gegensatz dazu steigt die kumulierte Inzidenz von Sekundärmalignomen, am ehesten verursacht durch die Bestrahlung, beginnend etwa 15 Jahre nach Lymphombehandlung exponentiell an und liegt nach 30 Jahren bei über 20 Prozent (Abb. 2).

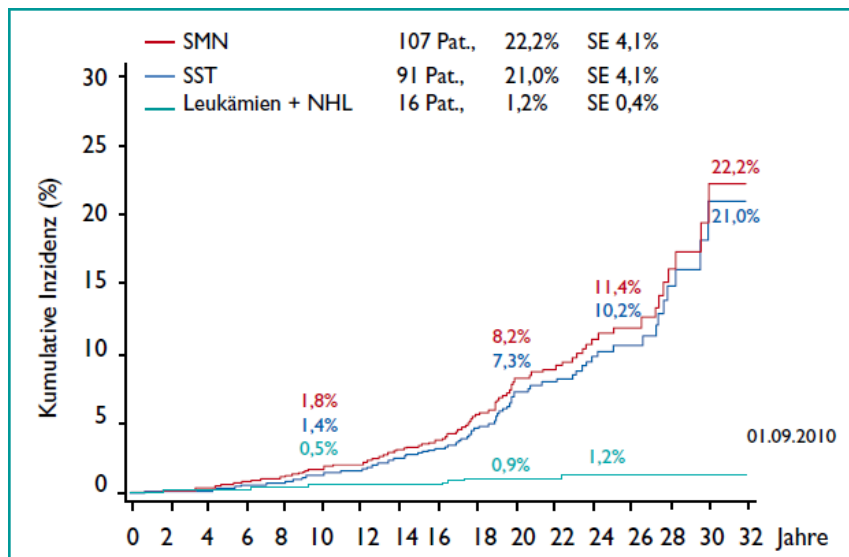


Abb. 2: Entwicklung von malignen Zweiterkrankungen nach Jahren der Therapie (Quelle: (Dörffel et al. 2010), SMN=sekundäre maligne Neoplasien, SST=sekundäre solide Tumoren)

Aufgrund der Seltenheit maligner Erkrankungen im Kindes- und Jugendalter (ca. 150 neu diagnostizierte HL-Erkrankungen pro Jahr in Deutschland) werden die Patienten nahezu lückenlos im Rahmen von zentralen, teilweise auch europaweit angelegten Therapieoptimierungsstudien (TOS) behandelt (Kurch et al. 2013b). TOS können so innerhalb kurzer Zeit eine hohe Zahl von Patienten rekrutieren, die eine standardisierte Therapie nach dem jeweiligen Stand der Wissenschaft erhalten. Zudem können auf schnellerem Weg neue Erkenntnisse gewonnen werden, die wiederum die nachfolgenden Behandlungsprotokolle beeinflussen. Die beiden Hauptziele sind bei allen TOS die Erhaltung bzw. die Verbesserung des Überlebens und bei Patienten mit sehr guten Heilungschancen die Reduktion therapieinduzierter Spätschäden (Dörffel et al. 2010); (Tallen et al. 2010). Zur Behandlung von Kindern und Jugendlichen mit einem Hodgkin-Lymphom werden seit 1978 TOS durchgeführt. Diese Doktorarbeit entstand im Rahmen der europaweit angelegten EuroNet-PHL-C1-Studie, welche im Zeitraum von 2007 bis 2013 unter Beteiligung von 18 Ländern durchgeführt wurde. Wesentliche Erkenntnisse der EuroNet-PHL-C1-, aber auch ihrer Vorgängerstudie (GPOH-HD-2002) wurden in die seit Oktober 2015 rekrutierende EuroNet-PHL-C2-Studie, die zehnte aufeinanderfolgende TOS, für das pädiatrische HL in Deutschland integriert (Mauz-Korholz et al. 2010); (Purz et al. 2011), (Hasenclever et al. 2014).

### **3.4 Ermittlung von Risikofaktoren zur Therapietitration und besondere Bedeutung der FDG-PET/CT**

Zum Erreichen hoher Heilungsraten bei optimal niedrigen Nebenwirkungsraten ist eine Therapietitration anhand bekannter, aber auch noch zu erforschender Risikofaktoren wichtig. Risikofaktoren bei Kindern und Jugendlichen mit Hodgkin-Lymphom sind E-Befall, d. h. Hodgkin-Beteiligung von extralymphatischen Geweben, ein Tumorbulk (Tumordurchmesser > 10 cm nach Lugano-Kriterien bzw. Tumolvolumen > 200 ml nach Definition der EuroNet-PHL-Gruppe) und eine B-Symptomatik. Der wichtigste, multivariat (d. h. bei gleichzeitiger Analyse verschiedener Variablen) unabhängige Risikofaktor ist das frühe Therapieansprechen nach zwei Kursen OEPA-Chemotherapie (Vincristin, Etoposid, Prednisolon, Adriamycin) in der FDG-PET/CT (Kurch 2011); (Kluge und Körholz 2011). So konnte (Kurch 2011) zeigen, dass eine Reduktion der glykolytischen Stoffwechselrate um mehr als 57 Prozent mit einem signifikant verbesserten ereignisfreien Überleben einhergeht. Das unterstreicht die Bedeutung der PET zur Beurteilung des Therapieansprechens. Auch (Lopci et al. 2011) zeigte bei einer kleinen Kohorte von 98 pädiatrischen Patienten aus acht italienischen Zentren, dass die FDG-PET der einzige unabhängige Prädiktor für das Gesamt- und ereignisfreie Überleben war.

Andere Autoren verweisen auf weitere Vorteile der PET in der Diagnostik und Therapie des Hodgkin-Lymphoms. (Robertson et al. 2011) beschrieben die besondere Bedeutung des PET für die Planung der Bestrahlungstherapie sowohl in Hinsicht auf die Definition des Zielvolumens als auch auf das Bestrahlungsfeld. (Paulino et al. 2012) untersuchten 53 Patienten. Bei neun von diesen musste das Bestrahlungsfeld durch die zusätzliche PET-Information angepasst werden. (Metwally et al. 2011) konnten nachweisen, dass bei Verwendung der PET/CT zur Bestrahlungsplanung die Interobservervariabilität deutlich geringer ausfällt als bei alleiniger Verwendung von CT-Bildern.

Denkbar ist auch, dass die Ausdehnung des durch die FDG-PET/CT diagnostizierten Skelettbefalls ein prognostischer Risikofaktor beim pädiatrischen Hodgkin-Lymphom ist. Hierzu bedarf es aber einer momentan nicht verfügbaren Methode, welche den Skelettbefall zu quantifizieren vermag.

### **3.5 Organisationsstruktur der EuroNet-PHL-C1-Studie und Einsatz der PET/CT**

Im Folgenden soll auf die Abfolge der bildgebenden, vor allem nuklearmedizinischen Verfahren im Rahmen der EuroNet-PHL-C1-Studie eingegangen werden. Die FDG-PET bzw. FDG-PET/CT wurde jeweils zum initialen Staging und zur Frühresponsebeurteilung (Early Response Assessment = ERA) nach Verabreichung von 2 Zyklen OEPA-Chemotherapie durchgeführt.

Die Bildgebung zum initialen Staging (neben FDG-PET auch morphologische Bildgebung wie Computertomographie, Magnetresonanztomographie und Sonographie) ist für die Therapieplanung wichtig, um anhand des Tumorstadiums die Intensität der Chemotherapie festzulegen und die Größe der Bestrahlungsfelder im Falle einer notwendigen Strahlenbehandlung zu bestimmen. Die ERA-PET dient der Evaluation des frühen Ansprechens auf die Chemotherapie und der Abschätzung der Prognose. Auf Basis der ERA-PET wird seit 2007 erstmalig weltweit beim pädiatrischen Hodgkin-Lymphom (PHL) auf die sich normalerweise an die Chemotherapie anschließende Radiotherapie verzichtet, wenn die ERA-PET-Untersuchung keinen pathologisch gesteigerten Glukosestoffwechsel aufweist (Kluge und Körholz 2011).

Alle lokalen Zentren führen zunächst selbst eine Begutachtung ihrer bildgebenden Verfahren durch. Zum Zwecke einer standardisierten Bewertung der bildgebenden Befunde erfolgt eine zweite, zentrale Referenzbegutachtung. Hierbei werden die PET, PET/CT- und PET/MRT-Bilder durch die Klinik für Nuklearmedizin des Universitätsklinikums Leipzig und die MRT- und CT-Bilder durch die Radiologie des Universitätsklinikums Halle erneut bewertet. Anschließend werden alle Befunde in einem interdisziplinären Tumorboard besprochen und die Ergebnisse an die lokalen Zentren übermittelt. Die Speicherung aller Bilddatensätze erfolgt auf einem zentralen, durch ein Multifirewallsystem geschützten Server, auf den nur autorisierte Personen per Passwort zugreifen können. Der Datenzugriff ist lediglich der Studienzentrale sowie den Referenzradiologen und Referenznuklearmedizinern erlaubt. Umgekehrt haben aber sämtliche teilnehmende Studienpartner die Möglichkeit, Bilddaten direkt an den zentralen Server zu senden, was den Versand der Bildgebung per CD oder DVD ersetzt. Zusätzlich bietet der zentrale Server die Möglichkeit, alle Bilddaten mit einem einheitlichen Programm auszuwerten und zu vergleichen, sodass ein



gesonderter Download entfällt. Die regelmäßige Wartung dieses Servers erfolgt in enger Zusammenarbeit mit der Firma Hermes Medical Solutions, Stockholm (Schweden). Im Rahmen einer EU-Förderung konnten über 130 Zentren in 18 europäischen Ländern im Zeitraum von 2008 bis 2011 angeschlossen werden (Kurch et al. 2013a).

Seit Beginn der EuroNet-PHL-C1-Studie im Jahre 2007 sind über 8000 PET- bzw. PET/CT-Datensätze auf den zentralen Server transferiert und referenzbegutachtet worden.

### **3.6 Detektion von Skelettbefall**

Ein Skelettbefall erfolgt beim Hodgkin-Lymphom klassischerweise hämatogen, ist charakterisiert durch ein uni- bzw. multifokales Befallsmuster und führt zu einem Upstaging in das Ann-Arbor-Stadium IV (Olson und Donaldson 2008). Im Rahmen der EuroNet-PHL-C1-Studie wird außerdem noch eine Unterteilung zwischen Knochenmarkbefall und Kortikalisbefall vorgenommen, da Läsionen mit einer Beteiligung der kompakten Knochensubstanz, nicht aber auf das Knochenmark begrenzte Läsionen in die Bestrahlungsfelder integriert werden (Kurch et al. 2013b). Von einem hämatogenen Skelettbefall abzugrenzen ist ein per-continuitatem-Befall durch in das Skelett einwachsende Lymphommassen. Dieser führt nicht zu einer Einstufung in das Stadium IV nach Ann Arbor, sondern wird als E-Läsion gewertet.

Die zur Verfügung stehenden bildgebenden Verfahren sind hinsichtlich ihrer Detektionsrate eines Skelettbefalls unterschiedlich sensitiv und spezifisch.

(Even-Sapir 2005) untersuchten hierzu verschiedene bildgebende Techniken wie CT, MRT, PET, Skelettszintigrafie und Röntgen:

Konventionelles Röntgen nimmt eine eher untergeordnete Rolle ein, da eine Dichtereduktion des Knochens von 30 bis 75 Prozent erforderlich ist, ehe ein Befall sichtbar wird (Even-Sapir 2005).

Die Skelettszintigrafie detektiert sehr sensitiv Bereiche erhöhter Osteoblastenaktivität und ist damit eine indirekte Methode zur Diagnose eines Knochenbefalls. Bereits ein Ersatz von 5 bis 10 Prozent physiologischen Knochens durch Tumor kann detektiert werden (Blake et al. 2001). Die Skelettszintigrafie weist für verschiedene Tumorentitäten eine erhebliche Varianz von Sensitivitäten im Bereich von 62 bis 100 Prozent und Spezifitäten im Bereich von 78 bis 100 Prozent auf (Hamaoka et al. 2004). Die Sensitivität kann noch um ca. 20 bis 50 Prozent bei Einsatz der SPECT bei

Wirbelsäulenmetastasen gesteigert werden (Gates 1998). Hingegen gelingt mittels Skelettszintigrafie kaum die sensitive Detektion einer Knochenmarksinfiltration, sodass das Ausmaß eines Knochenmarkbefalls nicht abgeschätzt werden kann (Evans et al. 2011). Bei heutzutage verfügbarer PET/CT spielt die Skelettszintigrafie zur Diagnose eines Knochenmarkbefalls beim Hodgkin-Lymphom nur eine untergeordnete Rolle (Martin 2009).

Mit der CT kann die Kortikalis hinsichtlich Kortikalisdestruktionen im Sinne lytischer Veränderungen bzw. die Knochenstruktur hinsichtlich mehrsklerosierter Areale untersucht werden. Zur Detektion von malignen Wirbelsäulenläsionen erreicht sie aber nur eine Sensitivität von 66,2 Prozent bei allerdings hoher Spezifität von 99,3 Prozent (Buhmann et al. 2009).

Die MRT hingegen stellt als weitere morphologisch-funktionelle Methode eine sehr sensitive Modalität zur Beurteilung des Knochenmarks auf Knochenmarksinfiltrationen dar und verursacht zudem keine Strahlenbelastung, was besonders bei Kindern und Jugendlichen von Bedeutung ist. So konnten (Buhmann et al. 2009) eine Sensitivität von 98,5 Prozent bei einer Spezifität von 98,9 Prozent bei malignen Wirbelsäulenläsionen verschiedenster Tumorentitäten nachweisen. Bei der Detektion von Knochenmarkbefall bei Lymphomen zeigte sich bei aggressiven Typen mit 88,9 Prozent eine deutlich höhere Sensitivität als bei indolenten Lymphomen mit 23,5 Prozent (Adams et al. 2013). Nachteilig ist die lange Untersuchungszeit sowie bei sehr kleinen Kindern die Erfordernis einer zusätzlichen Narkose und vor allem das meist begrenzte Untersuchungsfeld.

(Adams et al. 2014) zeigten in einer Metaanalyse bei 955 Patienten mit neu diagnostiziertem Hodgkin-Lymphom eine Sensitivität der FDG-PET/CT von 96,9 Prozent bei einer Spezifität von 99,7 Prozent zur Detektion von Knochenmarkbefall.

Da auch gutartige Veränderungen am Knochen (benigne Knochtumoren, Entzündungen, Morbus Scheuermann oder traumatische Verletzungen) zu Anreicherungen von  $^{18}\text{F}$ -FDG führen können, unterstützt die zusätzliche CT-Komponente bei der differenzialdiagnostischen Abwägung (Shulkin et al. 2009).

Eine weitere Schwierigkeit bei der Beurteilung der FDG-PET ist die diffuse Knochenmarkaktivierung, welche Ausdruck einer paraneoplastischen Reaktion ist und unterschiedlich stark ausgeprägt sein kann. Diese führt, wie unten stehende Abbildung 3 zeigt, zu einer Aktivitätsmehrbelegung der rotes Knochenmark enthaltenden Skelettregionen. (Elstrom et al. 2004) stellen fest, dass Knochenmarkaktivierung unter

Umständen schwierig von disseminiertem Skelettbefall zu differenzieren ist, was allerdings eher für die Entitäten des Non-Hodgkin-Lymphoms, nicht aber für die Entitäten des Hodgkin-Lymphoms zuzutreffen scheint. Abbildung 4 zeigt zum Vergleich einen Patienten mit fokalem Befall im Bereich des thorakolumbalen Übergangs ohne Knochenmarkaktivierung. Außerdem können auch Wachstumsfugen physiologische Mehranreicherungen aufweisen (Stauss et al. 2008), wie in Abbildung 5 gezeigt.

Insgesamt hat sich die Hybriduntersuchung FDG-PET/CT als die sensitivste und spezifischste bildgebende Methode zur Detektion von Skelettbefall bei Patienten mit Hodgkin-Lymphom herausgestellt.



Abb. 3: Patient mit Knochenmarkaktivierung



Abb. 4: Patient mit Läsionen in Wirbelsäule ohne Knochenmarkaktivierung

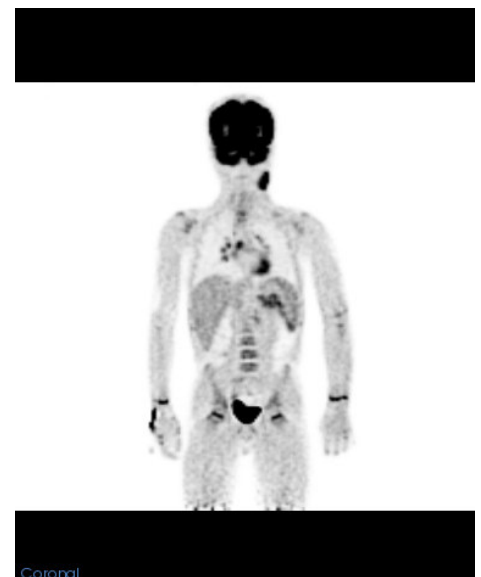


Abb. 5: Patient mit starker Aktivitätsbelegung der Epiphysenfugen der distalen Radii und proximalen Femora

### **3.7 Bedeutung von Computerprogrammen für die bildgebende Diagnostik, insbesondere für die Detektion eines Skelettbefalls**

Computerprogramme haben heute in der überwiegend volldigitalisierten bildgebenden Diagnostik eine herausragende Bedeutung. Sie ermöglichen z. B. in den Schnittbildverfahren über die Anfertigung multiplanarer Rekonstruktionen die exakte Beurteilung dreidimensionaler Bilddatensätze. Die Anwendung der Volume Rendering Technique (VRT) unterstützt die Visualisierung komplexer Zusammenhänge im 3D-Bild, z. B. die Lagebeziehung eines Tumors zu umgebenden Strukturen.

Computerprogramme können den befundenden Arzt aber auch als Detektionshilfe pathologischer Strukturen unterstützen. (Liang et al. 2016) untersuchten retrospektiv mit vier verschiedenen Programmen pulmonale Noduli in Low-Dose-CT-Untersuchungen von 50 Patienten im Rahmen des Lungenkrebsscreenings, die sich in einer Folgeuntersuchung als Lungenkrebs erwiesen. Hierbei detektierten die Programme in bis zu 70 Prozent der Fälle im initialen CT kleine Knotenbefunde, die zwar abgebildet, aber von den Befundern nicht diagnostiziert wurden und sich in der Folgeuntersuchung als Lungenkrebs herausstellten. Gleichzeitig wurden aber 20 Prozent der diagnostizierten Karzinome nicht durch die Programme erkannt. (Liang et al. 2016) unterstreichen deshalb die Rolle solcher Programme zur Zweitbefundung.

Neben der Detektion ist häufig auch die Charakterisierung des Ausmaßes einer Pathologie für die Prognose und Therapieabschätzung bedeutsam. In verschiedenen Bereichen der Medizin rücken nichtinvasive Methoden zur Diagnosesicherung in den Vordergrund. Zur Diagnosestellung, z. B. einer kardialen Amyloidose, sind eine Biopsie und histologische Analyse derzeitiger Goldstandard. Allerdings existiert mit der  $^{11}\text{C}$ -PIB-PET auch eine nichtinvasive Methode, um Amyloid-Ablagerungen in Herz und Gehirn abzubilden. Um diese aber in die klinische Praxis einzuführen, ist eine schnelle und reproduzierbare Methode notwendig. (Kero et al. 2015) stellen ein Programm vor, welches sich zur Quantifizierung der Amyloidplaques und Verlaufskontrolle nach Therapie eignet.

Auch in der Therapiesteuerung, z. B. im Rahmen einer Radiotherapie, sind Computerprogramme hilfreich. Hierbei ist die Abgrenzung der vitalen Tumorzellen als Zielvolumina mittels funktioneller Bildgebung durch PET für die Therapieplanung von besonderer Bedeutung und sollte standardisiert erfolgen. (Arens et al. 2014) testeten drei verschiedene Algorithmen bei 40 Patienten mit Hals- und Kopftumoren, die eine  $^{18}\text{F}$ -Fluorthymidin-(FLT)-PET/CT als Staging-Untersuchung erhielten. Hierbei erwies sich der FLAB (Fuzzy Locally Adaptive Bayesian)-Algorithmus als der zuverlässigste in der Segmentierung von Tumervolumina. Dieser Algorithmus bestimmt in einem vorgegebenen Volumen die Wahrscheinlichkeit, inwieweit jeder enthaltene Voxel Tumorgewebe versus physiologisches Gewebe repräsentiert, anhand der

- 1) Intensität des Voxels,
- 2) statistischen Verteilung der Voxel in diesem Volumen,

### 3) räumlichen Beziehung dieses Voxels zu den angrenzenden Voxeln.

Segmentierung hat in der Hybridbildgebung eine besondere Bedeutung, da anatomische und funktionelle Informationen gemeinsam betrachtet werden müssen. Wie (Bagci et al. 2012) betonen, muss diese besonders genau erfolgen, da die SUV-Berechnung durch die Regionengröße wesentlich beeinflusst wird. Die manuelle Abgrenzung wird aber stets durch Interobservervariabilität verfälscht. (Bagci et al. 2012) stellen daher einen „Random-Walk-Algorithmus“ vor, der PET- und CT-Bilder durch Informationen beider Methoden automatisch segmentiert. Bei einem „Random Walk“ (etwa „zufällige Irrfahrt“) als Sonderfall einer Markov-Kette erfolgt die Auswahl der Voxel nach einem zufälligen Muster, wobei stets nur der aktuelle Zustand für den nächsten Schritt entscheidend ist und jeder mögliche Nachbarpunkt im Raum mit der gleichen Wahrscheinlichkeit erreicht werden kann.

Für die computergestützte Analyse von Skelettläsionen gibt es bereits Konzepte in der zweidimensionalen Bildgebung. So hat die Firma Exini Diagnostics eine Software entwickelt, welche es ermöglicht, eine atlasbasierte Segmentierung von Skelettszintigrafiebildern vorzunehmen. Der im Programm hinterlegte Atlas wurde anhand von 734 Trainingsbildern erstellt. Die anschließende Segmentierung war in einer von (Sjostrand et al. 2009) veröffentlichten Studie bei 106 von 113 Testbildern erfolgreich. Im Quantifizierungsmodul werden der Bone Scan Index (BSI) und die Zahl der Metastasen kalkuliert.

Die Segmentierung in dreidimensionalen Schnittbilduntersuchungen ist ungleich komplizierter. Knochenspezifische Tracer wie  $^{18}\text{F}$ -Fluorid-PET reichern sich nur im Skelett an. (Puri et al. 2012) untersuchten bei vier postmenopausalen Frauen mittels  $^{18}\text{F}$ -Fluorid-PET/CT den Stoffwechsel der Hüfte, um das Frakturrisiko vorherzusagen. Sie unterteilten die Hüfte semiautomatisch in verschiedene Regionen und verwendeten aus der CT berechnete ROIs zur PET-Analyse. Sie betonen die Schnelligkeit und reduzierte Interobservervariabilität ihrer Methode gegenüber der manuellen Abgrenzung.

Die Analyse und Segmentierung des Skeletts in einer  $^{18}\text{F}$ -FDG-PET-Studie ist hingegen nicht ohne Weiteres möglich, da die Aufnahme von Glukose nicht spezifisch im Skelett erfolgt.

Allerdings ist es einer italienischen Arbeitsgruppe gelungen, das gesamte Skelett aus einer FDG-PET/CT-Untersuchung zu extrahieren. (Sambuceti et al. 2012) verwendeten die Datensätze von 102 Patienten mit nicht metastasiertem Melanom. Sie entwickelten einen Algorithmus, der es erlaubt, nicht nur das intraossäre Volumen, sondern auch das Volumen des roten Knochenmarks zu bestimmen und zu analysieren. Dieses Programm basiert auf der Annahme, dass intraossäres Volumen von Kompakta umgeben ist und diese den höchsten Hounsfieldwert aller Gewebe zeigt. Es wird ein Algorithmus angewandt, welcher die äußere Kortikaliskontur für jede einzelne CT-Schicht anhand ihrer Dichte erfasst. Um die nächste CT-Schicht zu analysieren, wird der optimierte Knochenumriss der vorhergehenden Schicht als neue Startkontur für den Algorithmus genutzt. Der Benutzer muss drei Schritte im Programm selbst vornehmen.

1. Anwahl einer Startschicht in der Schädelbasis
2. Überprüfung der Erfassung des Spinalkanals als extraossärer Raum
3. Entfernung von Kalzifikationen in extraskelettalem Gewebe

Im Übrigen verläuft dieses Programm vollautomatisch mit einer allerdings beträchtlichen Rechenzeit von zwei Stunden für jeden Patienten.

Die vorliegende Arbeit baut nun auf einem Programm auf, welches ebenfalls die Extraktion des Skeletts aus koregistrierten FDG-PET/CT-Untersuchungen erlaubt. Über die Dichtewerte in der CT-Komponente wird eine Segmentierung des Skeletts vorgenommen. Zur Diskrimination von Skelett und nichtskelettalen Geweben wird ein experimentell gefundener Wert von 160 Hounsfield Units (HU) als Grenzwert verwendet (Elsner et al. 2012). Gewebe, die diesen Cut-Off überschreiten, werden in einer sogenannten CT-Skelettmaske zusammengefasst. Diese Maske wird anschließend auf die PET-Komponente übertragen (siehe auch Methodikteil Kapitel 5.1.2).

Bisher existiert keine Methode, skelettale Herdbefunde in einer PET/CT-Untersuchung mittels eines Computerprogrammes zu detektieren.

### 3.8 Nicht-computergestützte, visuelle Bewertung des Skelettsystems im Rahmen der EuroNet-PHL-C1-Studie

In der EuroNet-PHL-C1-Studie werden neben 26 Lymphknotenstationen auch fünf Organe bzw. Organsysteme, unter anderem das Skelettsystem, beurteilt. Bei der Bewertung des Skeletts werden alle verfügbaren Bildgebungen herangezogen (in der Regel PET/CT, CT, MRT, seltener auch Skelettszintigrafie), insbesondere um Knochenmark- und Knochenbefall differenzieren zu können. Für den Skelettbefund in der PET stehen vier Items zur Klassifikation zur Verfügung:

- „0“ Kein Skelettbefall
- „1“ Skelettbefall
- „2“ Unklarer Befund
- „4“ Knochenmarkaktivierung

Ist mit einem der oben genannten bildgebenden Verfahren ein Skelettbefall detektiert worden, wird ein Dokument angelegt, in welchem detailliert die befallenen Regionen aufgelistet werden (siehe Abbildung 6).

GPOH-HD-2003		Knochenbefallsbogen				KB		
Patienten-ID: <input type="text"/>	Klinik-ID: <input type="text"/> - <input type="text"/> <small>Länderkürzel Kliniknummer</small>	Name, Vorname: <input type="text"/>		Geburtsdatum: <input type="text"/>				
Einrichtung: <input type="text"/>		<input type="radio"/> Staging <input type="radio"/> Restaging Nr. <input type="text"/> <input type="radio"/> Rezidiv-Staging <input type="radio"/> Rezidiv-Restaging Nr. <input type="text"/> <input type="radio"/> außerplanmäßig						
Untersuchungsdatum: <input type="text"/> PET: <input type="text"/>		Skelettszintigrafie: <input type="text"/>		CT: <input type="text"/>		MRT: <input type="text"/>		
		Region / Lokalisation:		PET	Knochen- szintigrafie	CT	MRT	Resümé
		1 - Kopf						
		2 - Wirbelsäule						
		3 - Becken						
		4 - Rippen						
		5 - Arm						
6 - Bein								
<b>Kodierung – allgemein:</b> 0 = kein Befall, 1 = Befall, 2 = positiv, aber nicht i. S. von Lymphom (Trauma) <b>Kodierung – Resümé:</b> 1 = Knochenmarkbefall 2 = Knochenbefall im Strahlenfeld 3 = Knochenbefall außerhalb des Strahlenfeldes		<b>Tumorkonferenz am:</b> <input type="text"/> Datum: <input type="text"/> Name (in Druckbuchstaben): <input type="text"/> Unterschrift: <input type="text"/>						
<small>           Studienleitung: Universität Leipzig, Kinderklinik – Prof. Dr. D. Köhler    ☎ +49 341 9726 151    ☎ +49 341 97 26 199    ✉ hodgkin@medizin.uni-leipzig.de            Studienkoordination: Universität Leipzig, Kinderklinik – PD Dr. C. Mauz-Körholz    ☎ +49 341 9726 201    ☎ +49 341 97 26 199    ✉ hodgkin@medizin.uni-leipzig.de            Datenmanagement: Koordinierungszentrum f. Klinische Studien Leipzig – Studiensekretariat    ☎ +49 341 9726 616 / 626    ☎ +49 341 97 25 629    ✉ studsek@kksl.uni-leipzig.de         </small>								

Abb. 6: Knochenbefallsbogen

Dieses Dokument unterteilt sechs skelettale Regionen – Schädel, Wirbelsäule, Becken, Rippen, Arme, Beine. Jede Skelettläsion wird der jeweiligen Region zugeordnet und mit den Bewertungen in den anderen verfügbaren bildgebenden Verfahren verglichen. Speziell für Läsionen in der FDG-PET erfolgte eine zusätzliche Deskription der gefundenen Aktivitätsanreicherung als fokal (eindeutig) oder diffus (fraglich).

Abbildung 7 zeigt hierzu exemplarisch im vierten Lendenwirbelkörper eine diffuse und nur geringgradige Aktivitätsmehrbelegung. Bezüglich der Bewertung dieser Auffälligkeit bestand eine hohe Interobservervariabilität. Im Gegensatz dazu stellten sich die deutlichen fokalen Mehranreicherungen im ersten Lendenwirbelkörper und zwölften Brustwirbelkörper für alle Untersucher eindeutig als Skelettbefall dar (Abb. 7).

Das Skelettsystem galt im EuroNet-PHL-C1-Protokoll als befallen, wenn

- eine Knochenmarkbiopsie ein positives Ergebnis ergeben hatte
- oder
- eine Region in der PET detektiert und in einer weiteren Bildgebung (CT, MRT oder Skelettszintigramm) bestätigt wurde (Kluge und Körholz 2011).

Im Rahmen der vorliegenden Dissertationsschrift galt die im Dokument zur Erfassung des Skelettbefalls dokumentierte Bewertung des Skelettsystems als Referenz für die semiautomatische Läsionendetektion durch ein entsprechendes Computerprogramm. Darüber hinaus wurden auch Läsionen berücksichtigt, die zwar nur in der PET detektiert wurden ohne Korrelat in einer weiteren Bildgebung, aber vom befundenden Nuklearmediziner als fokal (eindeutig) dokumentiert wurden („PET-only“-Läsionen).





Abb. 7: Eindeutige Läsionen in LWK 1 und BWK 12, fragliche Läsion in LWK 4

### **3.9 Bisherige Erkenntnisse zur Diagnostik von Skelettbefall mittels PET/CT beim pädiatrischen HL**

Das bereits erwähnte uni- bzw. multifokale Befallsmuster des Skeletts beim Hodgkin-Lymphom erklärt die geringe Sensitivität der Knochenmarkstanze. Knochen(mark)befall kann bei einer uni- oder bilateralen Knochenmarkbiopsie, bei der nur ein sehr kleiner Ausschnitt des Knochenmarks untersucht wird, in vielen Fällen nicht detektiert werden (Pelosi et al. 2008); (Cheng et al. 2011). (Cheng et al. 2011) untersuchten bei einem Kollektiv aus 31 Kindern, Adoleszenten und jungen Erwachsenen im Alter von 6 bis 24 Jahren mit Hodgkin-Lymphom die Ergebnisse von Knochenmarkstanze und PET/CT im Hinblick auf Skelettbefall. Sie konnten zeigen, dass eine Konkordanz von positivem Ergebnis in der PET/CT und der Knochenmarkbiopsie nur bei Befall der Beckenkämme (als Punktionsorte bei der Knochenmarkstanze) vorlag. Dadurch war der Ertrag einer invasiven und schmerzhaften Knochenmarkbiopsie mit sechs Prozent gering im Vergleich zu einer Konkordanz von 92 Prozent bei einer Weichgewebsbiopsie (z. B. bei einer

Lymphknotenbiopsie). Die PET/CT jedoch erlaubt als Ganzkörperuntersuchung eine Beurteilung des gesamten Skeletts. (Pelosi et al. 2008) und (Cheng et al. 2011) schlagen deshalb die PET-gesteuerte Knochenmarkbiopsie zur histologischen Sicherung vor.

Beim pädiatrischen Hodgkin-Lymphom evaluierten (Purz et al. 2011) die Genauigkeit der FDG-PET anhand von 175 Patienten. Die Schwierigkeit hierbei bestand darin, dass die Verifikation eines positiven PET-Befundes aus ethisch nachvollziehbaren Gründen nicht über eine histologische Sicherung erfolgen konnte, sodass Sensitivität und Spezifität der PET anhand von Ersatzmarkern abgeschätzt werden mussten. (Purz et al. 2011) verwendeten hierzu verschiedene Referenzlevel mit Surrogatmarkern. Eine Skelettregion wurde im Referenzlevel C als befallen gewertet, wenn eine positive Knochenmarkbiopsie vorlag, ein multifokales Befallsmuster bestand oder die Skelettläsionen auf die Chemotherapie ansprachen. Unter Annahme dieses Referenzlevels wurden eine Spezifität von 98,5 Prozent und ein positiver prädiktiver Wert von 95,6 Prozent beim Einsatz der PET zur Detektion von Knochenmarkbefall erreicht. Bei 25 der 45 PET-positiven Patienten war zusätzlich eine MRT vorhanden, wobei sich bei 23 dieser 25 Patienten eine positive Konkordanz zeigte. Hieraus ließ sich eine Sensitivität der FDG-PET von 92 Prozent ableiten. Von 23 vorhandenen CT-Untersuchungen waren 10 Untersuchungen positiv hinsichtlich eines Skelettbefalls.

(Purz et al. 2011) zeigten außerdem, dass Skelettbefall beim pädiatrischen Hodgkin-Lymphom ein typisches Befallsmuster des Stammskeletts (Wirbelsäule, Thoraxskelett, Becken, proximale Enden von Humeri und Femora) aufweist. Bei 32 von 45 Patienten mit Skelettbefall waren mindestens drei Herde im Stammskelett vorhanden. Im Rahmen von degenerativen Wirbelsäulenerkrankungen kommt es bei älteren Patienten zu Entzündungsreaktionen, insbesondere in der Lendenwirbelsäule, welche mittels PET/CT angezeigt werden können (Rosen et al. 2006). Derartige Veränderungen sind bei Kindern und Jugendlichen höchst selten, sodass falsch-positive Skelettbefunde nur selten zu erwarten sind. Das oben beschriebene Befallsmuster ist bei pädiatrischen Patienten leichter zu erkennen, da diese in der Regel noch keine degenerativen Erscheinungen aufweisen. Die Diagnose eines Skelettbefalls wird in der aktuellen EuroNet-PHL-C2-Studie in der Regel durch die PET/CT gestellt, sofern dieses bildgebende Verfahren drei Herde im Stammskelett detektiert. Sollte nur eine solitäre PET-positive Skelettläsion zu detektieren sein,

empfehlen (Purz et al. 2011) eine gezielte Biopsie derselben, falls ein Skelettbefall die Therapieempfehlung verändern würde. Im Gegensatz zur Biopsie bietet die PET/CT außerdem als Ganzkörperuntersuchung die Möglichkeit, das Ausmaß des Knochenmarkbefalls zu beurteilen.

## **4 Aufgabenstellung**

Das Ziel der Dissertationsschrift ist die Erarbeitung und Überprüfung eines computergestützten, semiautomatischen Algorithmus zur Detektion, Charakterisierung und Quantifizierung des Skelettbefalls bei Kindern und Jugendlichen mit einem Hodgkin-Lymphom. Hierbei sollen zunächst Probleme innerhalb der bereits bestehenden Programmversion herausgearbeitet und beschrieben werden. Im nächsten Schritt müssen Lösungen für diese Probleme gefunden und das Programm mithilfe dieser optimiert werden. Anschließend soll ein Schwellenwert ermittelt und getestet werden, welcher Mehranreicherungsherde im Skelett möglichst sicher als Lymphombefall identifiziert. Ziel ist es, die Benutzerfreundlichkeit des Programmes schließlich soweit zu erhöhen, dass der Einsatz sowohl in der klinischen Routine als auch in der Wissenschaft erfolgen kann. Im Rahmen der klinischen Routine soll es dem Arzt die Befundung erleichtern. Für die Wissenschaft ist die Quantifizierung des skelettalen Tumolvolumens interessant, da dieses ein möglicher Risikofaktor beim PHL, aber auch bei anderen Tumorerkrankungen darstellen kann.

Im Idealfall erlaubt die fortentwickelte Variante der Software die Detektion und Quantifizierung möglichst aller Befallsregionen in einer vordefinierten Skelettmaske, allerdings unter Ausschluss der oft nur geringgradig aktivitätsmehrbelegten und damit noch physiologischen Skelettanteile. Zudem soll das Programm möglichst zwischen befallenen Regionen und Mehranreicherungen durch Knochenmarkaktivierung bzw. Mehranreicherungen in Wachstumsfugen differenzieren können.

## 5 Materialien und Methoden

### 5.1 Stand des Computerprogrammes zu Beginn der Doktorarbeit und Optimierung seiner Funktionsweise

Das verwendete Programm (siehe Einleitung 3.7) erlaubt bereits mithilfe weniger Bearbeitungsschritte die Extraktion des Skeletts aus dem PET-Datensatz nach Übertragung einer über die spezifischen Dichtewerte von Knochen und Knochenmark erstellten CT-Skelettmaske. Zunächst wurden die Zuverlässigkeit dieser Skelettextraktionsfunktion überprüft, dabei auftretende Probleme identifiziert und anschließend Lösungen zur Optimierung erarbeitet und in das Programm integriert.

#### 5.1.1 Benutzeroberfläche

Sämtliche zur Bearbeitung der Aufgabenstellung erforderlichen PET/CT-Datensätze waren auf einem zentralen Server (Kurch et al. 2013) verfügbar. Die entsprechende Benutzeroberfläche, wie sie von der Firma Hermes Medical Solutions konzipiert wurde, ist nachstehend abgebildet (Abb. 8). Über die Patienten-Identifikationsnummer (ID) – in der Maske mit XXXX bezeichnet – konnten die gewünschten PET/CT-Studien gefunden und aufgerufen werden. Die weitere Bearbeitung erfolgte dann mit der bereitgestellten Software.

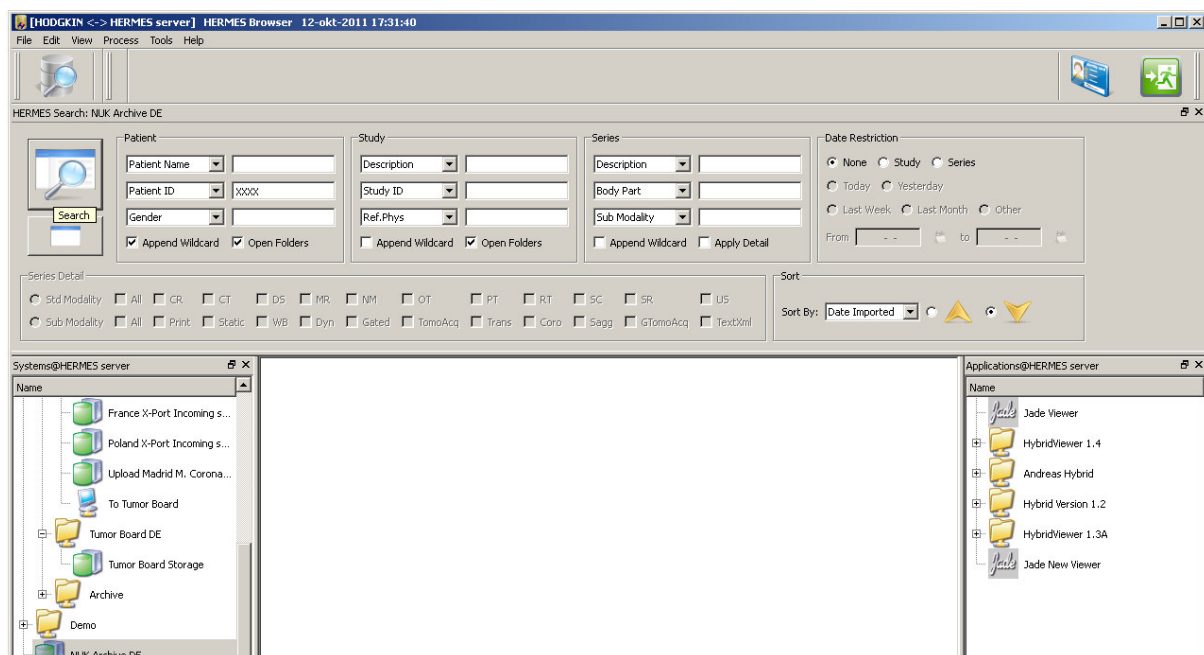


Abb. 8: Benutzeroberfläche des Hodgkin Super Servers mit Software Andreas Hybrid

### **5.1.2 Extraktion des Skelettsystems aus dem PET-Datensatz über eine Dichteschwelle der CT-Komponente und methodenimmanente Probleme**

Das PET-Bild selbst erlaubt eine visuell-qualitative Beurteilung des Skelettsystems sowie die Bestimmung semiquantitativer Parameter durch Platzierung einer Region of Interest (ROI) oder eines Volume of Interest (VOI) in dem interessierenden Bereich. Eine alleinig auf das Skelettsystem beschränkte, semiautomatische Analyse zur Anzeige verdächtiger Herdbefunde und deren gleichzeitige quantitative Charakterisierung sind allerdings bisher noch nicht möglich. Hierfür bedarf es eines Algorithmus, welcher die skelettale Nuklidbelegung von Aktivitätsanreicherungen des übrigen Ganzkörpers zu differenzieren vermag. Dieser bereits im Programm integrierte Algorithmus bedient sich der CT-Komponente aus der PET/CT: Die Hounsfielddichten, insbesondere von Kompaktasubstanz (ca. 250 - 1000 HU) und Knochenmark (ca. 30 bis 230 HU), weichen deutlich von denen anderer Gewebe (ca. -900 bis 90 HU) ab (Hofer 2010). Dieser Unterschied ermöglicht dem Programm zunächst innerhalb der CT durch Vorgabe eines HU-Cut-Off-Wertes, Knochensubstanz und Knochenmarkstrukturen bei Überschreiten dieses Schwellenwertes auszuwählen, wodurch die „CT-Skelettmaske“ erstellt wird, die wiederum mit dem PET-Datensatz überlagert werden kann. Die Überlagerung ermöglicht sodann die computergestützte Extraktion sämtlicher skelettaler Anreicherungen aus dem PET-Datensatz.

Vorarbeiten zu dieser Thematik haben einen Cut-Off-Wert von 160 HU zur vollständigen Erfassung des Skeletts ergeben (Elsner et al. 2012). Anteile des Knochenmarks, die eine geringere HU-Dichte als 160 aufweisen, werden allerdings nicht in der CT-Skelettmaske erfasst, an diesen Stellen finden sich dann Aussparungen („Holes“). Um dieses Problem zu beheben, wurde eine zusätzliche „Auffüllfunktion“ (s. g. „Hole Filling Filter“) implementiert. Dazu werden von allen nicht in die CT-Skelettmaske eingeschlossenen, aber direkt angrenzenden Voxeln zunächst die benachbarten Voxel bestimmt, die bereits zur Maske gehören. Sofern mehr als 50 Prozent der Nachbarvoxel von der CT-Skelettmaske erfasst werden, wird auch der bisher nicht berücksichtigte Voxel als Skelettvoxel bewertet und in die Maske integriert. Abbildung 9 zeigt in einer sagittalen Rekonstruktion (farblich rot gekennzeichnet), – dass vor allem im Bereich der Lendenwirbelsäule und des Sternums nicht alle Skelettstrukturen vollständig in der CT-Skelettmaske erfasst sind. Abbildung 10 zeigt

dasselbe Skelett nach Anwendung der Füllfunktion – sämtliche vormals rot markierten Anteile im Inneren der Wirbelkörper und des Sternums sind nunmehr in die CT-Skelettmaske eingeschlossen.



Abb. 9: CT-Skelettmaske vor Einsatz der Knochenfüllerfunktion mit Fehlstellen in Wirbelkörpern



Abb. 10: Knochenfüllerfunktion schließt Fehlstellen in CT-Skelettmaske

Nach Erstellung der CT-Skelettmaske erfolgt deren Projektion auf den PET-Datensatz (Abb. 11). Anschließend kann die skelettale Aktivitätsbelegung ohne Einschluss von Aktivitäten in den übrigen Geweben analysiert werden. Als problematisch für diesen Algorithmus erwies sich die gleichzeitige Integration von extraskelettalen Anteilen in die CT-Skelettmaske:

1. Sowohl orale als auch intravenöse Kontrastmittelapplikationen können durch die hohe Ordnungszahl des Kontrastmittels Dichtewerte vergleichbar mit Knochenkompakta und Knochenmark annehmen. Dadurch werden diese extraossären Depositionen fälschlicherweise in die CT-Skelettmaske integriert, wie unten stehende Abbildung 12 für im Darm befindliches Kontrastmittel zeigt. Diese extraskelettalen Depositionen können aber relativ einfach und vollständig unter visueller Kontrolle mithilfe einer in das Programm integrierten Ausschneidefunktion (s. g. „Crop VOI“) aus der CT-Skelettmaske entfernt werden. Hierzu wird der

extraskelettale Anteil manuell umfahren und kann in allen drei Ebenen betrachtet und ggf. angepasst werden. Anschließend kann der ausgewählte Anteil bestätigt und entfernt werden. Eine Entfernung ist auch dann möglich und nötig, wenn eine unmittelbare Nachbarschaft zum Skelett besteht.

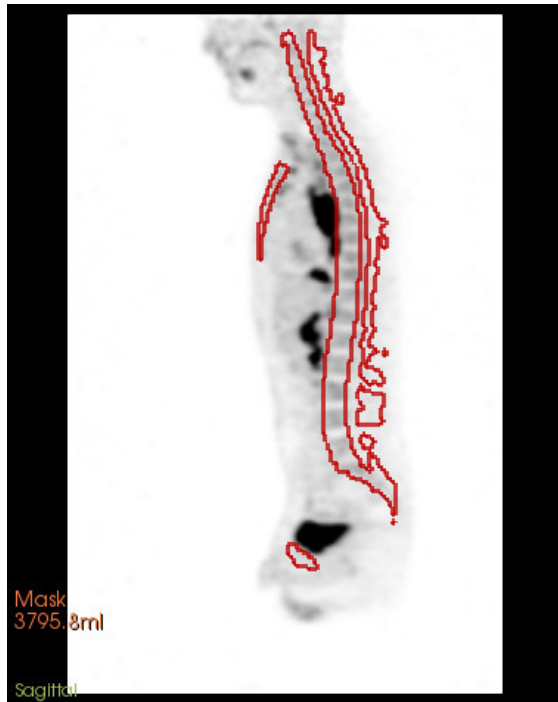


Abb. 11: CT-Skelettmaske wird auf PET übertragen

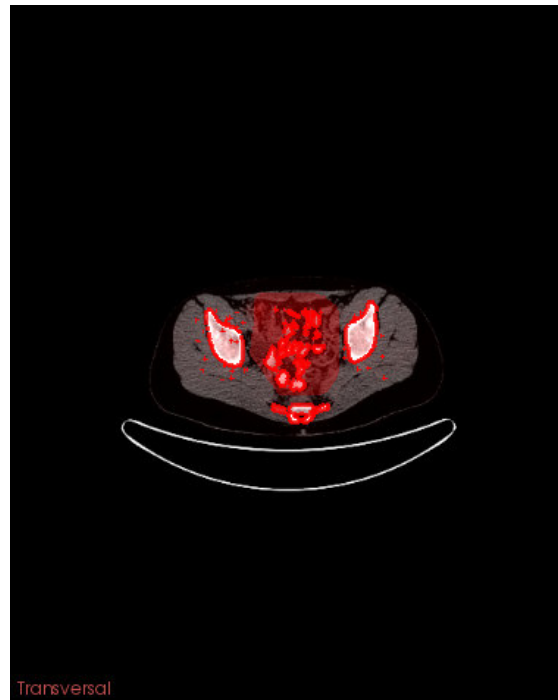


Abb. 12: Kontrastmittel im Darm wird mittels „Crop VOI“-Funktion markiert

Das eben beschriebene Problem betrifft prinzipiell alle mit oralem oder venösem Kontrastmittel gefüllten Hohlorgane wie z. B. Herz, große Gefäße, Niere, ableitende Harnwege (einschließlich Blase) sowie Dünn- und Dickdarm. Unten stehende Abbildung 13 zeigt in einer Sagittalrekonstruktion (mit roten Pfeilen markiert), wie bei einer Schwelle von 160 HU Teile der mit Kontrastmittel gefüllten Aorta und des Herzens miterfasst werden. Durch Anpassung der HU-Schwelle hin zu höheren Werten konnte dieser Effekt teilweise ausgeglichen werden. Wie in unten stehender Abbildung 14 erkennbar, nimmt die Einfärbung von Mediastinum, Harnblase und Darm mit zunehmender Hounsfieldddichte, hier exemplarisch bei 200 HU, ab.



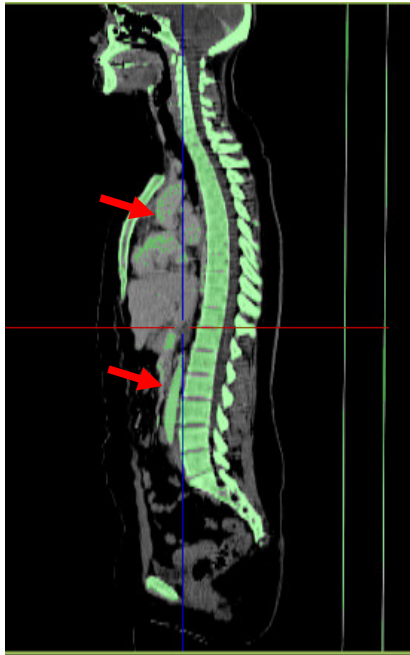


Abb. 13: CT-Skelettmaske bei 160 HU



Abb. 14: CT-Skelettmaske bei 200 HU

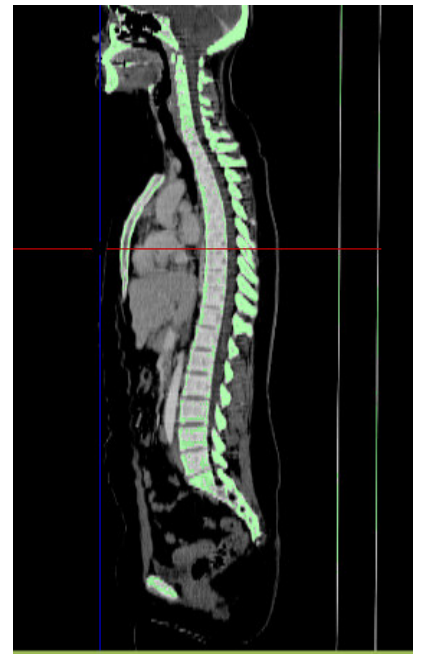


Abb. 15: CT-Skelettmaske bei 300 HU

Problematisch ist aber, dass die Knochenspongiosa z. T. Dichtewerte  $< 160$  HU (Hofer 2010) aufweist und deshalb durch Erhöhung der HU-Schwelle Teile des Skeletts nicht integriert werden. Dies wird exemplarisch in oben stehender Abbildung 15 deutlich: Bei einer Schwelle von 300 HU wird zwar einerseits kein extraskellettales, kontrastmittelhaltiges Gewebe mehr erfasst, andererseits werden große Teile des Skeletts nicht mehr markiert und bleiben somit für die spätere Skelettmaske unberücksichtigt. Diese Problematik muss im Rahmen der Optimierung des Programmes berücksichtigt werden. Zudem gilt es, eine Strategie zur optimalen Trennung von skelettalem und extraskellettalem Gewebe zu finden.

2. Die metallischen Anteile der Patientenliege werden aufgrund ihrer hohen Hounsfielddichte  $\gg 300$  regelhaft in die CT-Skelettmaske eingefasst. Bestandteile der Patientenliege gehören aber nicht zum Skelettsystem und müssen von diesem abgetrennt werden, da ansonsten falsche semiquantitative Kennwerte wie Durchschnitts- und Median-SUV für das gesamte Skelett daraus resultieren.

Die Möglichkeit der Extraktion des Skeletts aus der PET mithilfe einer CT-Skelettmaske wurde bei 142 pädiatrischen Hodgkin-Lymphompatienten der EuroNet-PHL-C1-Studie mit einem Skelettbefall überprüft. Hierbei galt es herauszufinden, welche HU-Schwelle bei der Mehrzahl der Patienten suffizient zur Anwendung kommen kann. Ferner testete man die bereits implementierten Funktionen (Hole-

Filling-Filter) und Werkzeuge (v. a. Ausschneidefunktionen) auf ihre Grenzen bzw. ihre Praktikabilität und erfasste den jeweils dafür notwendigen zeitlichen Arbeitsaufwand.

## **5.2 Experimentelle Ansätze zur softwaregestützten Detektion tumorbefallener Skelettanteile in der PET nach Programmoptimierung**

Die Möglichkeit zur Extraktion des Skeletts aus dem PET-Datensatz ist die Grundlage für die Ermittlung eines Algorithmus zur semiautomatischen Detektion von Tumorbefall im Skelett und zur quantitativen Bestimmung der skelettalen Tumorlast.

Da bisher keine Methode und kein semiquantitativer Grenzwert zur semiautomatischen Detektion von Skelettbefall bei Patienten mit Hodgkin-Lymphom existieren, wurden verschiedene, nachfolgend näher beschriebene Hypothesen aufgestellt und an ausgewählten Patientenkohorten getestet. Der sich dabei als am praktikabelsten herausstellende Ansatz soll an einer möglichst großen Zahl von pädiatrischen Hodgkin-Patienten mit Skelettbefall auf Robustheit hin überprüft werden. Die Auswertung der PET/CT-Datensätze erfolgte mit Genehmigung des Study Boards der EuroNet-PHL-C1-Studie.

### **5.2.1 Ansatz I: Analyse des Nuklidspeicherverhaltens im virtuell extrahierten Skelett durch Normierung auf das Speicherverhalten im Referenzorgan Leber**

Um Skelettbefall von physiologischem Skelettuptake geräteunabhängig (unterschiedliche PET-Geräte in einer Multizenterstudie) und standardisiert zu unterscheiden, kann die durchschnittliche und physiologische Impulsdichte im extrahierten Skelett mit der durchschnittlichen Impulsdichte einer extraskelletalen Referenzregion (z. B. Leber, Mediastinum) durch Bildung eines Quotienten ins Verhältnis gesetzt werden (Hasenclever et al. 2014). Anhand der Quotienten kann dann die statistische Berechnung eines Schwellenwertes erfolgen, welcher die Differenzierung von physiologischem Skelettuptake gegenüber pathologischer Mehrbelegung ermöglicht.

Als Skelettuptake wird in dieser Arbeit die durchschnittliche Aktivitätsbelegung innerhalb der CT-Skelettmaske definiert. Als extraskelletale Referenzregion sollte wiederum ein Organ verwendet werden, welches eine weitgehend homogene

Aktivitätsbelegung in einem ausreichend großen Volumen aufweist und zudem bei der Hodgkin-Erkrankung nur sehr selten befallen ist. Hierfür eignet sich die Leber im Besonderen. Sie zeigt einen relativ homogenen Uptake und ist bei weniger als fünf Prozent der PHL-Patienten befallen (Ilica et al. 2011).

(Hasenclever et al. 2014) haben eine entsprechende Methode bereits im Rahmen der PET-Frühresponseevaluation (ERA) entwickelt, bei der man die Leber als Referenzregion für die semiautomatische Bewertung der Restaktivität von Lymphomherden verwendete. Diese s. g. „qPET-Methode“ wurde an fast 900 Patienten getestet. Zur Ermittlung des qPET-Wertes wurde der Uptake der residuellen Lymphomherde ( $SUV_{peak}$  = Mittelwert der vier heißesten, unmittelbar zusammenhängenden Voxel) mit dem durchschnittlichen Uptake des Lebergewebes (mittlerer SUV-Wert einer 30 ml VOI) ins Verhältnis gesetzt. Von den ermittelten nahezu 750 qPET-Werten (bei 150 der 900 Patienten war kein zur Quantifizierung geeigneter Restuptake vorhanden) erstellte man Dichtekurven und analysierte deren Verlauf mathematisch. Mithilfe mathematischer Modelle konnten einmal ein sensitiver ( $qPET = 1,3$ ) und einmal ein spezifischer qPET-Cut-Off-Wert ( $qPET = 2,0$ ) berechnet werden, die jeweils zwischen normalverteilter und nicht normalverteilter Restglukosestoffwechselaktivität im Residuum und damit adäquater und nicht adäquater Frühresponse unterscheiden.

Analog hierzu wurde in der vorliegenden Arbeit an PET-Datensätzen von Patienten im Initialstaging eine quaderförmige VOI im rechten Leberlappen mit einem Volumen von 30 ml platziert, wie in Abbildung 16 gezeigt.



Abb. 16: Positionierung des 30 ml-Leberquaders

Um Kriterien zu erarbeiten, welche eine semiautomatische Erkennung des Skelettbefalls im abgebildeten Skelettsystem ermöglichen, untersuchte man **zufällig ausgewählte** Patienten im Initialstaging

- a) mit physiologischem Skelett,
- b) mit deutlichem Skelettbefall,
- c) mit Knochenmarkaktivierung sowie
- d) mit Skelettbefall zur Responsebeurteilung.

Keiner der untersuchten Patienten wies einen Leberbefall auf. Es wurden insgesamt 61 Datensätze analysiert. Die Zuordnung zu den Gruppen a bis c erfolgte entsprechend der Kriterien zur visuellen Bewertung des Skelettbefalls (siehe Kapitel 3.6). **21** Patienten wiesen **keine Beteiligung des Skeletts** und **keine Knochenmarkaktivierung** auf. Weiterhin wurden **20** Patienten mit **Skelettbefall** ausgewählt, die im Mittel 20,8 Herde [Range 5 bis 34] aufwiesen und bei denen mindestens fünf skelettale Herde einen Durchmesser von mehr als 1 cm hatten. Das Kriterium von mindestens fünf Herden mit jeweils mindestens 1 cm Durchmesser wurde unter der Annahme festgelegt, dass sich eine größere Tumorlast durch eine signifikante Erhöhung der Aktivitätsbelegung des befallenen Skeletts im Vergleich zum gesunden Skelett widerspiegeln müsste. Außerdem untersuchte man **zwölf** Patienten mit visuell geringgradiger (sieben Fälle) und ausgeprägter diffuser **Knochenmarkaktivierung** (fünf Fälle), um zu testen, ob diese Form der

Aktivitätsmehrbelegung von Skelettbefall unterscheidbar ist (bzgl. diffuser Knochenmarkaktivierung siehe Einleitung, Kapitel 3.6 mit Abb. 3).

Neben den eben genannten Patienten mit PET zum Initialstaging wurde zusätzlich auch eine kleine Gruppe von Patienten mit **Frühresponse-PET** untersucht (n=10). Diese Gruppe rekrutierte sich aus der o. g. Kohorte mit Skelettbefall (n=20, mindestens 5 Herde > 1cm) und zeigte bei initial ausgeprägtem Skelettbefall (im Mittel 20,8 Herde) ein deutliches Ansprechen der Skelettbefunde nach zwei Zyklen Chemotherapie (mit im Mittel nur noch einem Herd [Range 0-3]). Bei dieser Patientengruppe wurde erwartet, dass durch das Ansprechen auf die Chemotherapie und die daraus resultierende Abnahme der Aktivitätsbelegung der Tumorherde auch die durchschnittliche Aktivitätsbelegung des gesamten Skeletts verringert wird.

Für die vier Patienten-Gruppen wurden jeweils folgende Parameter erhoben:

Zwecks Bestimmung der Gesamtaktivitätsbelegung der CT-Skelettmaske erfasste man den Durchschnitts-SUV ( $SUV_{avg}$ ) mit Standardabweichung ( $SD_{suv}$ ) sowie den Median-SUV ( $SUV_{median}$ ) mit Range. Zur Charakterisierung der Aktivitätsbelegung des Lebergewebes wurden Durchschnitts-SUV ( $SUV_{avg}$ ), Maximal-SUV ( $SUV_{max}$ ) und Standardabweichung der Aktivitätsbelegung ( $SD_{suv}$ ) des 30 ml-Leberquaders ermittelt. Der Durchschnitts-SUV der CT-Skelettmaske und der Leber-VOI setzte man anschließend zueinander in Beziehung, um den qPET-Wert zu ermitteln.

### **5.2.2 Ansatz II: Histogrammanalyse**

Ein Histogramm ist eine grafische Darstellung der Häufigkeitsverteilung metrisch skalierten Merkmale. (<http://wirtschaftslexikon.gabler.de/Definition/histogramm.html>, zuletzt geprüft am 15.11.2018). Histogramme werden verwendet, um eine Häufigkeitsverteilung zu beurteilen und Spezifikationsgrenzen zu ziehen.

In der Annahme, dass sich die Countraten bei physiologischer Nuklidspeicherung im Skelett von den Countraten bei pathologischem Skelettuptake systematisch unterscheiden, wurde die Histogrammanalyse vorgenommen. Hierfür konnte die bereits auf der Auswerteplattform implementierte Applikation genutzt werden, die es ermöglicht, die Häufigkeitsverteilung der Countraten (Impulsdichte) in der PET-Komponente für alle Voxel im virtuell extrahierten Skelett darzustellen (Abb. 17). Die Abszisse bildet die Impulsdichte (Counts pro Sekunde) der Voxel und die Ordinate die Anzahl der Voxel ab. Im Anschluss an die Histogrammerstellung bestand die

Möglichkeit, sich bereits vorimplementierte Histogrammkenngrößen – s. g. „Analytics“ – anzeigen zu lassen.

Die elf aus der Arbeit von (Beare 2011) übernommenen „Analytics“ lauten:

Huang, Intermodos, Minimum, Isodata, Li, Maximum Entropy, Moments, Yen, Renyi Entropy, Shanbhag, Triangle.

Hierzu konnte über die vorinstallierte Schaltfläche „Analytics“ (Abb. 17) die Position der jeweiligen Kenngröße als orange gestrichelte Linie angezeigt werden. Die Positionen dieser Kenngrößen auf der Abszisse wurden aus dem Histogramm abgelesen und händisch übertragen.

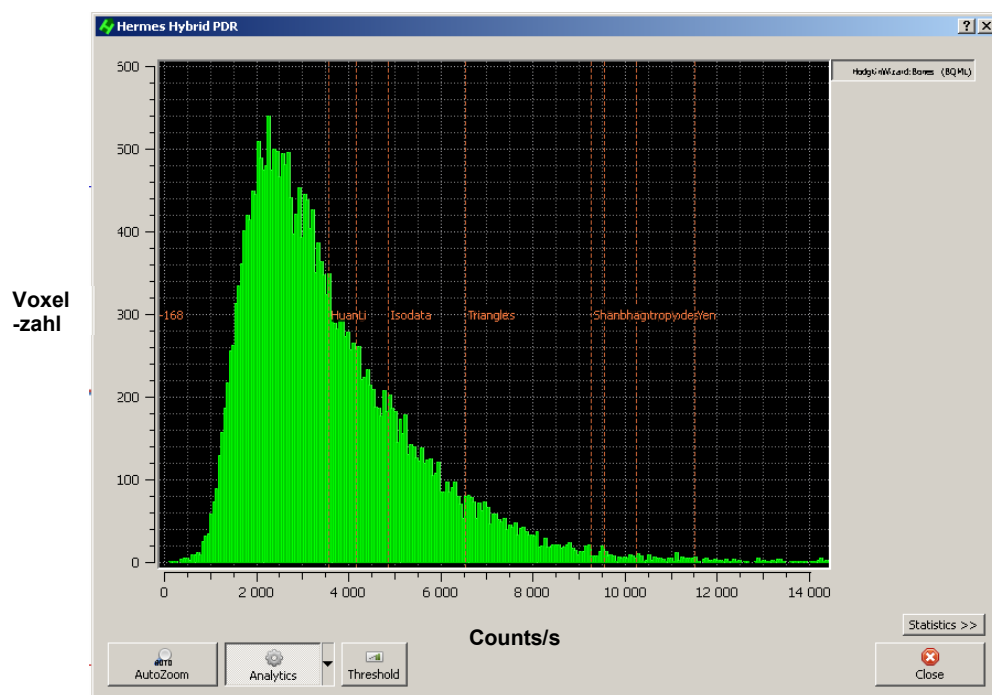


Abb. 17: Beispielhistogramm mit Kenngrößen

Diese Histogrammkenngrößen nutzen in ihrer Berechnung verschiedene mathematische Modelle, die z. B. beim Analytic Huang auf Shannons Entropiegleichung basieren. Shannons Entropiegleichung gibt dabei die durchschnittliche Mindestzahl von Bits an, die nötig sind, um eine Rangfolge von Zeichen zu kodieren, abhängig von Anzahl und Häufigkeit dieser Zeichen (siehe <http://www.shannonentropy.netmark.pl/>). Andere Analytics bilden nach iterativer Glättung der Histogrammkurven eine Schwelle zwischen zwei verbliebenen Peaks ab, wie z. B. Intermodos und Minimum (Beare 2011).

Die Idee hinter der Verwendung von Histogrammkenngrößen ist, eine Grenze zwischen physiologischer und pathologischer Anzahl an Skelettvoxeln zu finden bzw. diese mathematisch voneinander unterscheiden zu können.

Es bestanden somit insgesamt zwei Möglichkeiten der Histogrammanalyse:

Zum einen die rein visuelle Beurteilung der Histogrammkurven bei visuell physiologischer und pathologischer Nuklidbelegung des Skeletts, zum anderen die mathematische und schwellenwertbasierte Analyse.

Es wurden erneut die Datensätze der in Ansatz I (siehe Kapitel 5.2.1) zufällig ausgewählten Patienten hierfür herangezogen:

- a) die 21 Patienten ohne Auffälligkeiten im Skelettsystem,
- b) 15 der vormals 20 Patienten mit mindestens fünf abgrenzbaren Skelettläsionen > 1 cm und
- c) die zwölf Patienten mit einer diffusen Knochenmarkaktivierung unterschiedlichen Ausmaßes, wobei diese bei sieben Patienten gering bis mäßiggradig und bei fünf Patienten sehr ausgeprägt war.

Insgesamt erstellte man somit von 48 PET/CT-Datensätzen Histogramme, in denen die Impulsdichte (Counts pro Sekunde) der Voxel auf der Abszisse gegen die Anzahl der Voxel mit dieser Impulsdichte auf der Ordinate abgebildet wurde.

Eine erste Zwischenauswertung ergab folgendes Problem:

Die Histogramme der unten stehenden Abbildungen 18 und 19 zeigen exemplarisch zwei Patienten ohne Skelettbefall. Hierbei fällt auf, dass sich die Maximalwerte der Impulsdichte bei beiden Patienten deutlich unterscheiden (8000 Counts versus 16000 Counts) und dass die maximale Anzahl der Voxel mit einer bestimmten Aktivitätsbelegung ebenfalls differiert (maximal 250 versus 650).

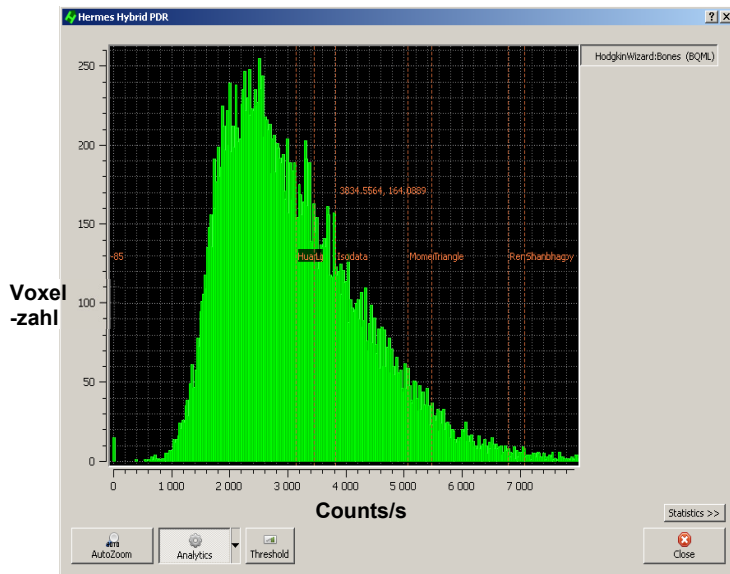


Abb. 18: Histogramm von Pat-ID 3202 ohne Skelettbefall

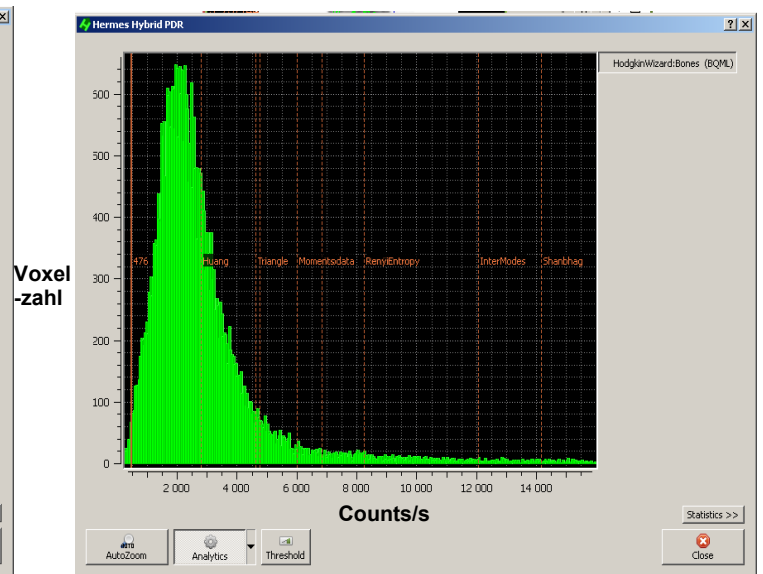


Abb. 19: Histogramm von Pat-ID 3161 ohne Skelettbefall

Damit sind die Zahlenwerte in den Histogrammen nicht vergleichbar, obwohl hinter beiden Histogrammen Patienten mit physiologischem Tracerbesatz des Skeletts stehen. Folglich waren zum Erreichen von vergleichbaren Werten verschiedene Schritte der Normierung erforderlich. Hierzu wurde zunächst der gewichtsbasierte  $SUV_{bw}$  (Abszisse) ins Programm implementiert. Zusätzlich erfolgte der Bezug der Aktivitätsbelegung auf das Referenzorgan Leber, um den Skelettbefall von physiologischem Skelettuptake geräteunabhängig und standardisiert zu unterscheiden, vgl. Kapitel 5.2.1 und (Hasenclever et al. 2014).

Des Weiteren war aus folgenden Gründen eine Normierung der Ordinatenachse an Stelle der Countrate erforderlich:

- 1) Die ausgewerteten PET/CT-Datensätze stammten aus einer multizentrischen Studie, sodass sich je nach Gerätetyp und angewandten Rekonstruktionsalgorithmen eine große Bandbreite der Voxelzahl/ml zwischen verschiedenen PET-Datensätzen ergab (siehe Ergebnisse 6.2.7).
- 2) Die Altersspanne der 48 Patienten reichte von 3 bis 18 Jahren mit den sich daraus ergebenden altersspezifischen Unterschieden des absoluten Skelettvolumens und der absoluten Voxelzahl des Skeletts.



Daher wurde eine Quotientenbildung aus dem Anteil der mit einer bestimmten Aktivitätsbelegung versehenen Voxel am Gesamtvolumen der CT-Skelettmaske vorgenommen, da durch diese relative Skala eine Unabhängigkeit von Voxelgröße und Rekonstruktionsalgorithmen wie auch der Körpergröße erzielt wird.

Unten stehende Abbildungen 20 und 21 zeigen die abschließend vom Programm generierten Histogramme der beiden Beispielpatienten ohne Skelettbefall. Auf der Abszisse wurde der Quotient aus der Aktivitätsbelegung der einzelnen Voxel dividiert durch die Aktivitätsbelegung des Leberquaders, quasi der qPET-Wert jedes einzelnen Voxels aufgetragen. Die Abszisse des Histogramms bezeichnete man als „rel. to background“ (bezogen auf den Hintergrund [d. h. der Leber]). Die Ordinate wird als „Volume (Prozent)“ bezeichnet und ist der Anteil aller Voxel mit einer bestimmten Aktivitätsbelegung an der Gesamtzahl der Voxel in der CT-Skelettmaske.

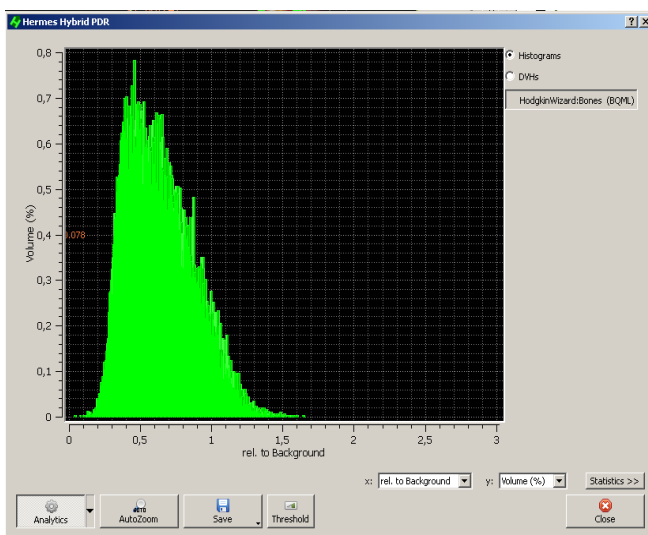


Abb. 20: Histogramm von Pat-ID 3202 ohne Skelettbefall

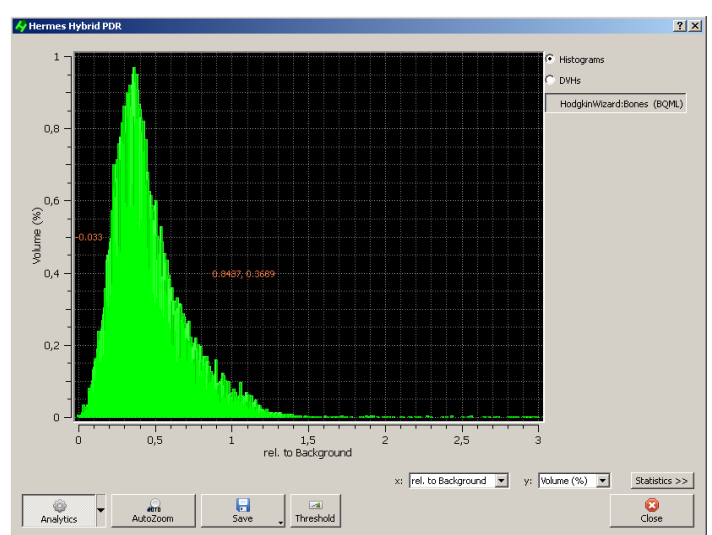


Abb. 21: Histogramm von Pat-ID 3161 ohne Skelettbefall

Die Skalenbreite der Histogramme kann vom Anwender flexibel angepasst werden. Es hat sich bei der Histogrammerstellung gezeigt, dass nur einzelne Voxel eine Aktivitätsbelegung von mehr als dem dreifachen der durchschnittlichen Leberaktivitätsbelegung aufwiesen. Daher wurde ein qPET-Wert von 3 als Grenze gewählt, um die Histogramme untereinander visuell vergleichen zu können.

### **5.2.3 Ansatz III: Aktivitätsanalyse ausgewählter Skelettregionen mit und ohne Lymphommanifestationen**

Da bekanntermaßen das Skelettsystem aus unterschiedlich dichten Knochenstrukturen besteht (zum Beispiel weist das Beckenskelett höhere Dichtewerte als das Rippenskelett auf), sollte geprüft werden, ob ein regionenspezifischer Cut-Off-Wert für verschiedene Skelettbereiche erforderlich ist und ob eventuell regionenspezifische Schwellenwerte zur Detektion von Skelettbefall verwendet werden müssen.

Hierzu wurde die Aktivitätsbelegung, getrennt nach Skelettregionen, bei 29 Patienten untersucht. Erneut wurden aus den in Ansatz I verwendeten Patienten zufällig elf Patienten mit ausgeprägtem Skelettbefall im Initialstaging (mindestens fünf abgrenzbare Skelettläsionen > 1 cm) sowie 18 im Skelett lymphomfreie Patienten ausgewählt.

Bei den Patienten mit Skelettbefall dokumentierte man 222 Läsionen im Rahmen der visuellen Referenzbegutachtung: 85 Läsionen befanden sich in der Wirbelsäule, zehn im Thoraxskelett (Rippen und Sternum), 101 im Becken, 9 im Schultergürtel (Oberarmknochen und Skapula) und 17 in den Oberschenkelknochen. Bei zehn dieser elf Patienten befundete man wie in Ansatz 1 das zusätzlich verfügbare Restaging zur Frühresponsebeurteilung, wodurch insgesamt 39 Datensätze bei 29 Patienten ausgewertet wurden.

Bei allen 39 PET/CT-Datensätzen fand eine Unterteilung des Skeletts in die sechs Regionen Wirbelsäule, Schädel, Rippenthorax, Becken, obere und untere Extremität statt. Jede dieser Regionen wurde unter visueller Kontrolle manuell eingegrenzt. Hierbei nutzte man erneut die bereits im Kapitel 5.1.2 beschriebene Ausschneidefunktion (Crop VOI), mit der manuell der zu untersuchende Bereich in der CT-Skelettmaske eingegrenzt wird. Im Resultat entstand eine Maske, die die zu fokussierende Skelettregion enthielt, wie exemplarisch für alle sechs Regionen unten stehende Abbildung 21 zeigt: In der ersten Reihe sind von links nach rechts Wirbelsäule, Becken, Rippenthorax inklusive Brustbein dargestellt, in der zweiten Reihe von links nach rechts Schädel, proximaler Oberschenkelknochen und proximaler Oberarmknochen mit Schulterblatt.

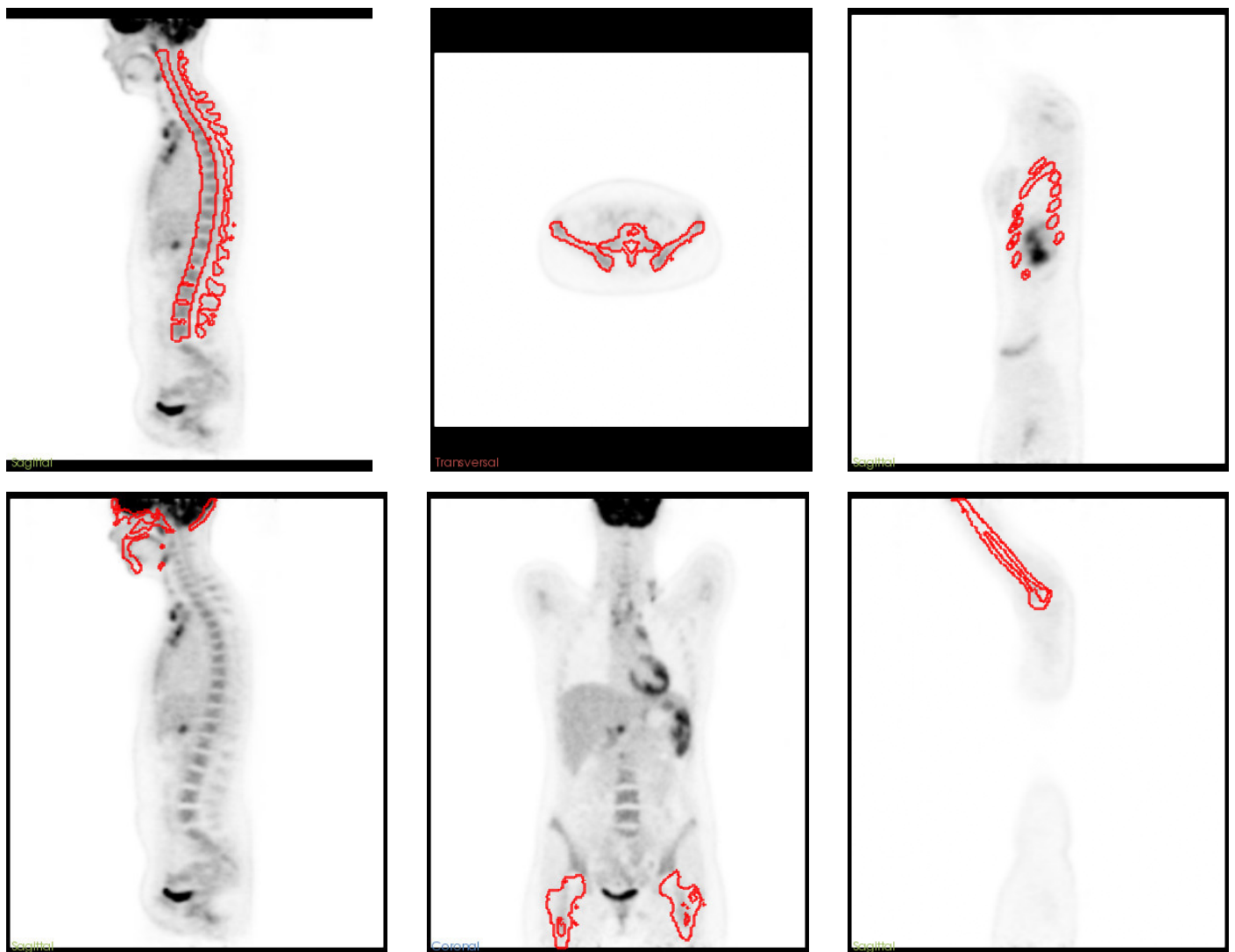


Abb. 22: Darstellung der sechs verschiedenen Skelettregionen zur Aktivitätsanalyse in unterschiedlichen Schnittebenen

Es wurden die Kennwerte  $SUV_{avg}$  und  $SUV_{max}$  für alle sechs Regionen erhoben und auf den  $SUV_{avg}$  eines 30 ml großen Leberquaders normiert:

Unter den Patienten mit physiologischem Skelett wiesen zwei Patienten einen deutlich erhöhten durchschnittlichen Leber- $SUV_{avg}$  bei 2,46 bzw. 2,24 und zwei Patienten einen deutlich verminderten Leber- $SUV_{avg}$  von 0,46 bzw. 0,87 auf. Diese Patienten schienen besonders geeignet, um erneut die Praktikabilität der Normierung auf die Leber zu testen.

#### **5.2.4 Ansatz IV: Vergleich der Aktivitätsbelegung von Wirbelkörpern und Wirbelkörperläsionen**

Während vorangegangene Arbeitsschritte vor allem die Normierung der Impulsdichte des Skeletts auf die Leber zum Gegenstand hatten, soll nun der direkte Vergleich zwischen physiologischer und pathologischer Nuklidbelegung des Skeletts bzw. von Skelettanteilen erfolgen. Diesem Ansatz liegt die Vorüberlegung zugrunde, dass das Skelettsystem in der FDG-PET einen physiologischen Nuklidbesatz aufweist. Bei einem Skelettbefall existiert somit ein Nebeneinander von physiologisch nuklidbelegten und pathologisch nuklidmehrbelegten Knochenstrukturen. Wenn das Skelett, wie in vorangegangenen Kapiteln beschrieben, mithilfe des CT-Datensatzes aus dem PET-Datensatz extrahiert wird, so muss zur Differenzierung zwischen physiologischem und pathologischem Nuklidbesatz ein Cut-Off-Wert gefunden werden, der nur Läsionen zählt, die diesen Cut-Off-Wert erreichen oder überschreiten. Ansonsten würden die physiologischen Nuklidanreicherungen bei der Gesamtergebnisermittlung berücksichtigt und dadurch zu einem fälschlicherweise niedrigeren Ergebnis der pathologischen Nuklidmehrbelegung führen.

In diesem Kapitel (Ansatz IV) wird nur die Wirbelsäule betrachtet, die beim Hodgkin-Lymphom die am häufigsten befallene Skelettregion darstellt. Im Folgekapitel 5.2.5 (Ansatz V) sollen dann alle Skelettläsionen im Hinblick darauf betrachtet und getestet werden, ob durch den Bezug auf eine gesunde skelettale Referenzregion die Detektion des Skelettbefalls gelingt.

##### **a) Physiologischer Uptake eines Wirbelkörpers als Referenz zur Unterscheidung zwischen physiologischem und pathologischem Skelettuptake**

Zunächst wurde untersucht, wie sich die Aktivitätsbelegung eines komplett befallenen Skelettabschnitts von einem sehr ähnlichen Abschnitt mit ausschließlich physiologischer Aktivitätsbelegung unterscheidet. Hierzu wurde die Aktivitätsbelegung eines nach PET-Kriterien gesunden Wirbelkörpers mit der Aktivitätsbelegung eines lymphombefallenen Wirbelkörpers verglichen. Zunächst wurde alleinig auf Wirbelkörper fokussiert, da durch den häufigen Befall der Wirbelkörper beim Hodgkin-Lymphom zeitgleich sowohl physiologische als auch pathologische, jedoch strukturell ähnliche Skelettabschnitte in enger räumlicher Nähe zueinander vorliegen. Diese

lassen sich aufgrund der gleichen Form und des ähnlichen Volumens sowohl visuell als auch semiquantitativ gut miteinander vergleichen. Es wurde ein möglichst großer, gut abgrenzbarer Wirbelkörper ohne Skelettbefall ausgewählt (s. g. Referenzwirbelkörper), um zu prüfen, ob durch Vergleich dessen physiologischer Aktivitätsbelegung mit der der übrigen Wirbelkörper diejenigen mit Läsionen (im Sinne von Lymphombefall) detektiert werden können.

Hier galt es zunächst, eine VOI zu konstruieren, welche den ganzen Wirbelkörper einfasst, um im Anschluss innerhalb der VOI Kennzahlen wie  $SUV_{peak}$  oder  $SUV_{mean}$  ermitteln zu können. Dies geschah programmgestützt durch das Werkzeug „RegionGrow“, welches von einem zentralen Startpunkt ausgehend alle Voxel einschließt, deren Aktivitätsbelegung einen bestimmten Basis-SUV überschreitet. Hierzu war die visuelle Kontrolle durch den Untersucher erforderlich. Nach Anklicken der Mitte des Wirbelkörpers wurde zunächst eine sehr wenige Voxel umfassende Start-VOI farblich markiert. Unter visueller Kontrolle konnte durch Bedienung der Maus die SUV-Schwelle erhöht und damit die anfänglich sehr kleine VOI weiter aufgezoogen und um Nachbarvoxel vergrößert werden, bis der gesamte Wirbelkörper erfasst war. Anschließend wurde das ebenfalls bereits im Programm implementierte Werkzeug „Metabolic Volume“ eingesetzt, dessen Algorithmus nach den vier Voxeln mit der höchsten Radiotraceraufnahme („4 Hottest Voxel“) in einer vorgegebenen VOI sucht. Dessen Algorithmus ist so konzipiert, dass zunächst der Voxel mit der höchsten Aktivitätsbelegung ( $SUV_{max}$ ) im vordefinierten Suchbereich, in diesem Fall der Wirbelkörper, bestimmt wird. Anschließend werden die drei unmittelbar an diesen angrenzenden Nachbarvoxel mit der nächsthöheren Aktivitätsbelegung ermittelt. Die mittlere Aktivitätsbelegung dieser vier Voxel wird automatisch bestimmt (sogenannter  $SUV_{peak}$ ). Vorteile des  $SUV_{peak}$  sind die hohe Reproduzierbarkeit und geringe Anfälligkeit für Partialvolumeneffekte im Vergleich zum  $SUV_{max}$  vgl. (Wahl et al. 2009). Zusätzlich zum  $SUV_{peak}$  wurden die durchschnittliche Aktivitätsbelegung ( $SUV_{avg}$ ) und die Standardabweichung ( $SD_{suv}$ ) zur Charakterisierung des Referenzwirbelkörpers erfasst.

## b) Befallsläsionen in Wirbelkörpern

Zwei Eigenschaften der Läsionen in jeweils befallenen Wirbelkörpern wurden bestimmt: Zum einen wurde das metabolische Volumen der Läsion erfasst, um eine Vorstellung zu erhalten, welche Größe Wirbelsäulenläsionen beim Hodgkin-Lymphom annehmen können. Zum anderen wurde der Durchschnittswert der Aktivitätsbelegung der vier Voxel mit der höchsten Aktivitätsbelegung in der Läsion als  $SUV_{peak}$  ermittelt. Als Startwert für die Volumenbestimmung in einer Wirbelkörperläsion wurde der im Vorfeld bestimmte  $SUV_{peak}$  des Referenzwirbelkörpers genutzt und als unterste SUV-Schwelle (Cut-Off) für die die Läsion umschließende VOI definiert. Unter visueller Kontrolle wurde die SUV-Schwelle so lange angepasst, bis die gesamte Läsion in die VOI inkludiert war. Der sich ergebende SUV-Grenzwert wurde als  $SUV_{search}$  dokumentiert. Die Vollständigkeit der Läsionserfassung wurde jeweils in allen drei Rekonstruktionsebenen überprüft.

Das sich innerhalb der VOI ergebende metabolische Läsionsvolumen (automatische Anzeige durch das Programm) wurde tabellarisch erfasst. In der so als Suchbereich definierten Läsionen-VOI wurde anschließend der oben beschriebene Algorithmus „4 Hottest Voxel“ eingesetzt, um den  $SUV_{peak}$  des Wirbelsäulenherdes zu erhalten. Anschließend wurde ein Quotient mit dem  $SUV_{peak}$  des physiologischen Wirbelkörpers gebildet.

Abbildung 23 soll beispielhaft die Methodik verdeutlichen: Zunächst wurde der nicht befallene, rot umrandete dritte Lendenwirbelkörper als Referenzregion über das Werkzeug „Region Grow“ erfasst und dessen  $SUV_{peak}$  bestimmt. Die anschließend (wie oben beschrieben) vermessenen Läsionen in Brust- und Lendenwirbelsäule sind mit zufällig vom Programm ausgewählten Farben markiert worden. In diesen Läsionen sind darauf jeweils die vier Voxel mit der höchsten Impulsrate als  $SUV_{peak}$  bestimmt und mit einer anderen Farbe markiert worden. Der  $SUV_{peak}$  wird in der rechts der sagittalen PET-Rekonstruktion stehenden Tabelle als avg.SUV (der betreffenden vier Voxel) angegeben.

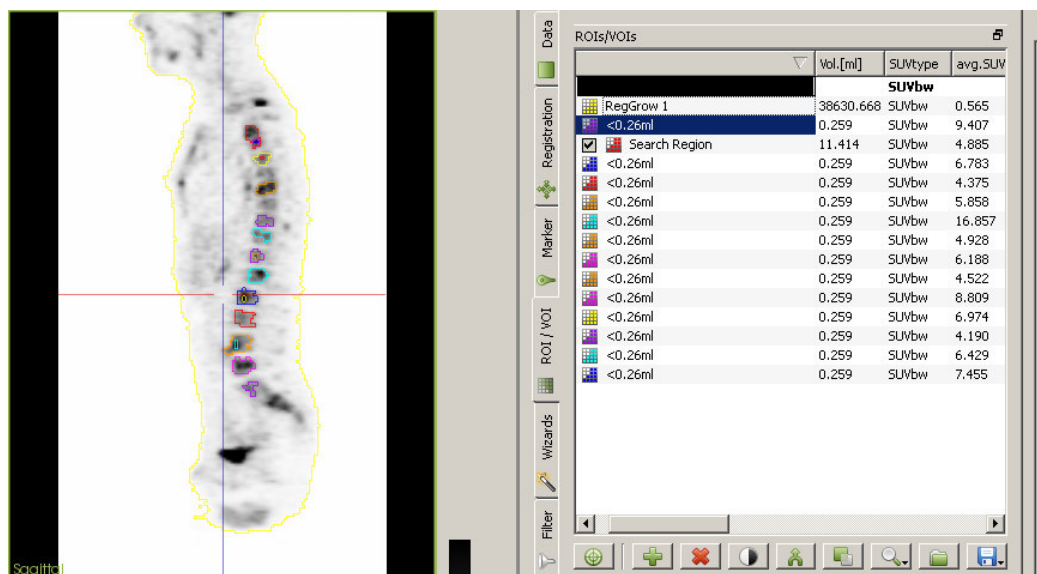


Abb. 23: Vermessung einzelner Läsionen und Ermittlung SUV<sub>peak</sub>

Auf die beschriebene Weise wurden die 20 Patienten mit deutlichem Skelettbefall aus Ansatz I (siehe Kapitel 5.2.1), die insgesamt 157 Wirbelkörperläsionen aufwiesen, untersucht. Es wurden für acht Läsionen in der Halswirbelsäule, 89 Läsionen in der Brustwirbelsäule und 60 Läsionen in der Lendenwirbelsäule die jeweiligen quantitativen Parameter bestimmt. Bei 20 Patienten mit jeweils sieben Hals-, zwölf Brust- und fünf Lendenwirbelkörpern ergeben sich potenzielle Befälle in 140 Hals-, 240 Brust- und 100 Lendenwirbelkörpern. In der hier untersuchten Patientenkohorte waren somit 8/140 (5,7 Prozent) der Halswirbelkörper, 89/240 (22,3 Prozent) der Brustwirbelkörper und 60/100 (60 Prozent) der Lendenwirbelkörper von einem Hodgkin-Lymphombefall betroffen (basierend auf dem Goldstandard der visuellen Referenzbegutachtung).

#### **5.2.5 Ansatz V: Test verschiedener Standardabweichungen über dem Mittelwert der Aktivitätsbelegung des Gesamtskeletts bzw. eines physiologischen Wirbelkörperuptakes als Grundlage für die programmgestützte Läsionendetektion**

Als vielversprechende Ansätze wurden die computergestützte Erfassung von Skelettläsionen in Bezug auf das gesamte Skelett und auf einen gesunden Wirbelkörper untersucht.

Bei der Erarbeitung von Ansatz IV (siehe Ergebnisse Kapitel 6.2.4 Unterpunkt b)) zeigte sich, dass einzelne Voxel in nicht befallenen Wirbelkörpern dennoch Aktivitätsbelegungen in der gleichen Höhe wie Lymphommanifestationen annehmen können, somit gibt es einen Überlappungsbereich zwischen physiologischen und pathologischen Aktivitätsbelegungen. Daher muss überprüft werden, ob zur Abtrennung einer pathologischen Läsion von gesundem Gewebe eine Mindestgröße von Voxeln festgelegt werden muss. Darüber hinaus sollte geprüft werden, ob durch Wahl eines Schwellenwertes mit  $n=x$  Standardabweichungen über der durchschnittlichen Aktivitätsbelegung einer Referenzregion die Abgrenzung zu Skelettläsionen möglich ist.

Denkbar wäre, dass dieser Schwellenwert nicht für alle Skelettregionen gleichermaßen anwendbar sein wird, da laut Auswertung von Ansatz III (siehe Ergebnisse Kapitel 6.2.3) einige Skelettregionen (Wirbelsäule und Becken) physiologischerweise höhere Aktivitätsbelegungen als andere (Rippenthorax und Extremitäten) aufweisen.

a) Semiautomatische Erfassung von Skelettläsionen unter Bezugnahme auf den Gesamtskelettuptake

Zur experimentellen Testung wurden erneut die Datensätze der elf Patienten aus Ansatz III verwendet, die im Referenzbogen für Skelettbefall mindestens fünf Skelettläsionen aufwiesen. Um die korrekte topografische Erfassung von Skelettherden zu den einzelnen Regionen zu gewährleisten, unterteilte man das Skelett in **45 Regionen**: Sternum, 24 Wirbelkörper, acht Einzelregionen im Becken (jeweils beidseits: Ossa ilii, Ossa ischii, Ossa pubis und Os sacrum), vier Regionen in der unteren proximalen Extremität (Femurkopf und proximaler Femur beidseits), vier Regionen in der oberen proximalen Extremität (Humeruskopf und proximaler Humerus beidseits), Rippenthorax beidseits und Schultergürtel beidseits. Die Aktivitätsbelegung der Läsionen wurde als fokal oder diffus bewertet (siehe Einleitung Kapitel 3.6).

Es wurde ein Suchalgorithmus verwendet, der alle Voxelgruppen (VOI) mit 1-, 2-, 2,5-, 3- und 3,5-facher Standardabweichung über dem Hintergrund der durchschnittlichen Aktivitätsbelegung des gesamten Skeletts detektiert, um zu prüfen, welcher dieser sich daraus ergebenden fünf Schwellenwerte die beste Konstellation hinsichtlich Sensitivität und Spezifität ergibt.



Es stellte sich heraus, dass Skelettanteile mit Knochenmarkaktivierung überproportional häufig als Läsionen angezeigt wurden. Dies betraf nicht das gesamte Skelett gleichermaßen, sondern vor allem Regionen mit einem hohen Anteil an rotem Knochenmark, wie die Wirbelsäule oder das Becken (vgl. Kapitel 5.2.4). Da unter Anwendung des Algorithmus bei diffuser Knochenmarkaktivierung für die beiden letzten Skelettregionen eine hohe Rate kaum herausfilterbarer, falsch-positiver Befunde zu erwarten ist, wurde dieser Weg nicht weiterverfolgt.

b) Semiautomatische Erfassung von Skelettläsionen unter Bezugnahme auf den Uptake eines nicht lymphombefallenen Wirbelkörpers

Wirbelkörper sind in der Regel von einer Knochenmarkaktivierung mitbetroffen, da sie bei Kindern einen hohen Anteil an rotem Knochenmark aufweisen (Elstrom et al. 2004). Wirbelkörper sind daher besonders geeignet, einen Cut-Off-Wert festzulegen, der über dem Niveau der Knochenmarkaktivierung liegt. Damit dürfte sich die Anzeige falsch-positiver Läsionen bei gleichzeitig vorliegender Knochenmarkaktivierung verringern. Gleichzeitig sollte aber die Detektion von Skelettläsionen, die im Regelfall einen höheren Uptake als die Knochenmarkaktivierung haben, gelingen.

Darüber hinaus zeigte sich in der Aktivitätsanalyse ausgewählter Skelettregionen mit und ohne Lymphommanifestation (siehe Kapitel 5.2.3), dass neben den Beckenknochen auch Wirbelkörper höhere physiologische Aktivitätsbelegungen als die übrigen Skelettabschnitte aufweisen. Daher ist zu erwarten, dass sich durch Verwendung eines Wirbelkörperuptakes als Referenz für den physiologischen Skelettuptake der Anteil falsch-positiv angezeigter Skelettläsionen deutlich verringert.

Wirbelkörper lassen sich durch die annähernd zylindrische Form gut in drei Raumebenen erfassen, und man hat bei der manuellen Eingrenzung im Gegensatz zu einer Analyse des gesamten Wirbels mit Pedikeln, Dorn- und Querfortsätzen weniger Partialvolumeneffekte. Daher wurde nun, anders als im Kapitel 5.2.4 a), der Wirbelkörper nicht über die „RegionGrow“-Funktion, ausgehend von den vier Voxeln mit der höchsten Aktivitätsbelegung, manuell eingegrenzt, sondern über die schneller anzuwendende, bereits in Kapitel 5.1.2 vorgestellte „Crop VOI“-Funktion.

Da Lendenwirbelkörper die größten Wirbelkörper sind, lassen sich diese besser als z. B. Halswirbelkörper manuell abgrenzen. Man verwendete vorzugsweise den vierten

Lendenwirbelkörper (LWK 4) als Referenz. Dieser steht regelhaft waagrecht zur Transversalebene und lässt sich gut in Koronalrekonstruktion und Sagittalrekonstruktion umfahren. Wenn jedoch LWK 4 einen Hodgkin-Befall aufwies, wurde der fünfte Lendenwirbelkörper verwendet oder, insofern auch dieser befallen war, der kranial davon am nächsten gelegene, nicht befallene Wirbelkörper als Referenz ausgewählt.

Abbildung 24 zeigt, wie der vierte Lendenwirbel zunächst in der Sagittalrekonstruktion eingegrenzt wird. Abbildung 25 zeigt, wie anschließend in der korrespondierenden Transversalrekonstruktion nur der Wirbelkörper ohne Wirbelbogen und Fortsätze umschlossen wird. Abbildung 26 bildet diesen Wirbelkörper in der Sagittalrekonstruktion erneut ab.



Abb. 24: Markierung des 4. Lendenwirbels mit „Crop VOI“-Funktion



Abb. 25: Körper des vierten Lendenwirbels wird in der Transversalrekonstruktion markiert



Abb. 26: LWK 4 als Referenzregion in der Sagittalrekonstruktion

Bei Skelettbefall von LWK 4 musste bei vier von elf Patienten auf einen benachbarten Lendenwirbelkörper und bei zwei von elf Patienten mit Befall der gesamten Lendenwirbelsäule auf einen Brustwirbelkörper ausgewichen werden.

Es wurden alle vom Programm detektierten VOIs mit ihren Volumina bei 2-facher, 2,5-facher und 3-facher Standardabweichung über dem  $SUV_{mean}$  des Wirbelkörpers tabellarisch erfasst.

Vorarbeiten haben gezeigt, dass die meisten Skelettläsionen mindestens vier zusammenhängende Voxel aufweisen (Georgi et al. 2018). Vom Algorithmus angezeigte Läsionen haben daher eine Mindestgröße von vier Voxeln.

VOIs, die eine Größe von ein bis drei Voxeln haben, werden vom Programm nicht als Läsionen angegeben, da auch einzelne Voxel in physiologischem Gewebe höhere Aktivitätsbelegungen aufweisen können (siehe auch Kapitel 5.2.4).

Es wurde die Anzahl falsch-positiver Läsionen zur Beurteilung der Spezifität erfasst.

### **5.2.6 Überprüfung des als am besten geeigneten Algorithmus an Patienten mit physiologischem Skelettuptake**

Nachdem der Suchalgorithmus mit 2,5 Standardabweichungen über der durchschnittlichen Aktivitätsbelegung eines Referenzwirbelkörpers das beste Verhältnis von Spezifität zu Sensitivität erbrachte, wurden 50 PET/CT von zufällig ausgewählten Patienten ohne Skelettbefall (analog Referenzbegutachtung) getestet, um die Spezifität des nachstehenden Algorithmus zu definieren:

1.)  $SUV_{Läsion} > SUV_{mean} \text{ des Wirbelkörpers} + 2,5 \text{ SD}$  und 2.)  $Größe_{Läsion} > 4 \text{ Voxel}$

Es wurden Anzahl, Lokalisation, metabolisches Volumen und  $SUV_{mean}$  falsch-positiver Läsionen sowie Durchschnitts-SUV und Standardabweichung des vierten Lendenwirbelkörpers tabellarisch erfasst.

### **5.2.7 Überprüfung und Optimierung des Größenkriteriums zur Anwendung auf Datensätze mit unterschiedlicher Matrixgröße**

Es sollte geprüft werden, ob ein absolutes Volumen oder eine absolute Voxelzahl als Schwellenwert der Größe für den Algorithmus (Schwellenwert der Aktivität:  $SUV_{Läsion} > SUV_{mean} \text{ des Wirbelkörpers} + 2,5 \text{ SD}$ ) geeigneter sind, da je nach Gerätetyp und angewandter Rekonstruktion eine große Bandbreite der Voxelzahl/ml zwischen verschiedenen PET-Datensätzen aufgefallen war (siehe Vorarbeiten zur Histogrammanalyse Kapitel 5.2.2). Somit ist denkbar, dass bei PET-Datensätzen mit sehr wenigen Voxeln/ml kleine Läsionen als negativ erkannt, hingegen bei PET-Datensätzen mit sehr vielen Voxel/ml häufiger Voxel mit eigentlich physiologischem Uptake als falsch-positive Läsionen durch das Programm angezeigt werden.

Es wurden die 142 Patienten mit Skelettbefall betrachtet, für die die Einschlusskriterien aus dem Folgekapitel 5.3, Unterpunkt 2 zutrafen. Es wurde aus den PET-Header-Daten ermittelt, wie viele Voxel in einem Milliliter Volumen enthalten sind. Diese Werte

wurden tabellarisch erfasst und verglichen. Wie die Auswertung (siehe Ergebnisse Kapitel 6.2.7) zeigte, ergaben bei 75 Patienten und damit mehr als der Hälfte 16 Voxel etwa 1 ml, vier Voxel entsprechen also 0,25 ml.

Daher wurde nun unter Anwendung des Algorithmus (mit dem Kriterium für die Aktivitätsbelegung der Läsion  $> \text{SUV}_{\text{mean}} \text{ des WK} + 2,5 \text{ SD}_{\text{WK}}$ ) untersucht, ob die Läsionen eher detektiert werden, wenn als Grenzgröße wie bisher vier Voxel oder ein Volumen von 0,25 ml verwendet wird. Aus den 142 zur Verfügung stehenden Patienten wurden für diese Analyse stichprobenhaft 36 Patienten ausgewählt. Diese wiesen insgesamt 256 Skelettläsionen auf.

### **5.3 Evaluation des experimentell gefundenen, besten Programmalgorithmus an einer großen Patientenpopulation zur Detektion und Quantifizierung des Skelettbefalls mittels PET**

#### **1.) Patientenkollektiv**

Es wurden alle verfügbaren pädiatrischen Patienten mit Hodgkin-Lymphom anhand der vorliegenden Patientenbögen seit Beginn der EuroNet-PHL-C1-Studie 2007 bis zum 31.11.2012 in die Auswertung einbezogen ( $n = \text{ca. } 2100$ ). Die Auswertung der PET/CT-Datensätze erfolgte mit Genehmigung des Study Boards der EuroNet-PHL-C1-Studie.

#### **2.) Einschlusskriterien**

Folgende Voraussetzungen für einen Patienteneinschluss mussten hierzu gegeben sein:

- a) In den Patientenunterlagen sollte im Referenzbogen für Evaluation von Skelettbefall (siehe Einleitung 3.8) mindestens eine PET-positive Läsion dokumentiert sein.
- b) Zum initialen Staging wurde eine schwächungskorrigierte PET/CT durchgeführt. Diese sollte die Anwendung des Suchalgorithmus zulassen, d. h.

sie sollte koregistriert sein, sodass PET- und CT-Datensatz kongruent aufeinander projizierbar sind.

- c) Durch Erfassung des Körpergewichts sollte die Berechnung von einem auf das Körpergewicht bezogenen SUV möglich sein.

### **Diese Voraussetzungen trafen bei 142 Patienten zu.**

#### **3.) Retrospektive vs. prospektive Testung**

Diese 142 Patienten wurden noch einmal in zwei Gruppen aufgeteilt:

- a) Bei 131 Patienten wurde retrospektiv untersucht, ob die visuell gefundenen und im Referenzbogen für Skelettbefall dokumentierten Läsionen im Skelett durch den semiautomatischen Algorithmus regelrecht detektiert wurden. Dabei wurde unter anderem überprüft, ob der bei zunächst elf Patienten als optimal erkannte Algorithmus (siehe Kapitel 5.2.5 b)) auf eine größere Kohorte anwendbar ist. Des Weiteren wurden die Sensitivität und Spezifität gegenüber der visuellen Referenzbegutachtung ermittelt (visuelle Referenzbegutachtung = Goldstandard).
- b) Weitere elf Patienten wurden prospektiv untersucht, wobei bei diesen nur das Vorliegen von Skelettbefall bekannt war, nicht jedoch die jeweiligen Lokalisationen. Erst nach Anwendung des Algorithmus verglich man die vom Programm erfassten Läsionen mit der visuellen Referenzbefundung. Hierdurch sollte der Einfluss der Subjektivität des Anwenders reduziert werden.

#### **4.) Ablauf des Algorithmus**

- a) Die CT-Skelettmaske erfuhr manuell eine Anpassung dergestalt, dass alle extraskelletalen Bezirke mit hohen HU entfernt wurden (siehe Kapitel 5.1.2). Normalerweise wird bei der PET/CT-Untersuchung eines PHL-

Patienten ein größerer Teilkörperabschnitt aufgenommen, der von der Schädelbasis bis zur Mitte des Oberschenkels reicht. In der Regel werden dabei die Arme eleviert neben dem Kopf positioniert, wodurch oft nur der proximale Humerus miterfasst wird. Diese Lagerung dient dazu, Fehler in der Schwächungskorrektur und Strahlenaufhärtingsartefakte zu vermeiden (vgl. (Barrett und Keat 2004)). Einigen Patienten ist die elevierte Armhaltung nicht möglich, beispielsweise wegen Lymphommanifestationen im Hals-/Schulterbereich, sodass die Arme seitlich am Körper anliegend aufgenommen werden. Außerdem existieren unterschiedliche Aufnahmestrategien, von Schädelbasis bis Mitte Oberschenkel und von Scheitel bis Mitte Oberschenkel, sodass mitunter Teile des Schädelknochens nicht zur Abbildung kommen. Bei Patienten mit ausgedehntem Skelettbefall werden in der Regel noch die unteren Extremitäten mit aufgenommen, um die dort selten vorkommenden Lymphommanifestationen zu detektieren. Bei den hier untersuchten 142 Patienten betraf das drei Patienten mit Läsionen distal von Humerus- und Femurschaftmitte. Da diese nicht in der standardisierten CT-Skelettmaske mit erfasst werden, erfolgte die Bestimmung der Läsionsgröße unter Anwendung der „RegionGrow“-Funktion (siehe Kapitel 5.2.4) mit den Parametern des Algorithmus. Das ermittelte Volumen wurde zu den übrigen Skelettläsionen addiert, um eine korrekte Quantifizierung des Skelettbefalls zu gewährleisten.

Um aber untersuchen zu können, inwieweit die Ausdehnung des Skelettbefalls als Prognosefaktor fungieren kann, muss die CT-Skelettmaske verschiedener Patienten aus den immer gleichen Skelettanteilen bestehen. Diese reicht von der oberen Halswirbelsäule bzw. den proximalen Oberarmknochen bis zu den proximalen Oberschenkelknochen und muss vom Anwender mithilfe der „Crop VOI“-Funktion (siehe Kapitel 5.1.2) zugeschnitten werden.

Eine so zugeschnittene, auf den PET-Datensatz übertragene Skelettmaske ist in den Abbildungen 27 und 28 in Sagittal- und Koronalrekonstruktion dargestellt.



Abb. 27: Sagittalrekonstruktion einer standardisierten, auf den PET-Datensatz übertragenen Skelettmaske



Abb. 28: Koronalrekonstruktion einer standardisierten, auf den PET-Datensatz übertragenen Skelettmaske

- b) Bei allen Patienten wurde ein nach visueller Einschätzung nicht befallener Wirbelkörper als Referenzregion ausgewählt und, wie unter Kapitel 5.2.5 b) erläutert, manuell eingegrenzt. Um den Referenzwirbelkörper genau zu charakterisieren, wurden seine Lokalisation, durchschnittliche Impulsrate und Standardabweichung der Impulsrate tabellarisch erfasst.
- c) Der Suchalgorithmus detektierte skelettale Herde, wenn folgende Bedingungen erfüllt waren:

1.  $SUV_{Läsion} > SUV_{mean} \text{ des Wirbelkörpers} + 2,5 \text{ SD}$  und
2.  $Größe_{Läsion} > 0,25 \text{ ml}$

## 5.) Erfasste Parameter

- a) Bei allen Patienten wurde nachträglich visuell auf Basis der semiautomatischen Befunde zwischen hämatogenem Skelettbefall und Skelettbefall durch Tumorwachstum per continuitatem (E-Befall)

unterschieden. Diese beiden Befallsmuster können nicht vom Programm unterschieden werden, da nur Voxel in der CT-Skelettmaske analysiert werden und nicht, ob ein größerer Anteil der Läsion außerhalb dieser Maske liegt, wie es bei Tumorwachstum per continuitatem der Fall wäre.

- b) Die durchschnittliche Impulsrate des Wirbelkörpers mit physiologischer Nuklidspeicherung soll auf ihre Eignung als Surrogatmarker zur Detektion des Ausmaßes der Knochenmarkaktivierung untersucht werden. (Murata et al. 2006) wiesen eine Korrelation der FDG-Belegung des Knochenmarks mit der Leukozytenzahl und insbesondere der Neutrophilenzahl nach. (Murata et al. 2006) argumentieren, dass die Neutrophilen mit mehr als 50 Prozent die dominierende Zellreihe des hämatopoetischen Knochenmarks sind und deren Lebensspanne mit zehn Stunden so kurz ist, dass die Neutrophilenzahl den Zustand der Granulopoese abbildet.
- c) Um die Sensitivität des Algorithmus zu berechnen, wurde die Anzahl der in der visuellen Referenzbegutachtung als Skelettläsionen befundeten Mehranreicherungen (Grundlage Referenzbogen für die Evaluation von Skelettbefall, siehe Abbildung 6 und Kapitel 3.8 der Einleitung) mit der Anzahl der vom Programm detektierten Läsionen verglichen. Zur weiteren Charakterisierung der Skelettläsionen wurden der Durchschnitts-SUV und das metabolische Volumen all dieser Läsionen erhoben. Nicht vom Programm detektierte, aber im Referenzbogen als Skelettbefall dokumentierte Läsionen untersuchte man dahingehend, ob die Nichterfassung auf eine zu geringe Aktivitätsbelegung bzw. Größe im Vergleich zu den Suchparametern des Algorithmus zurückzuführen ist oder ob diese evtl. gar nicht von der CT-Skelettmaske erfasst worden sind.
- d) Um zu überprüfen, ob bestimmte Regionen besonders häufig befallen sind, wurde die Lokalisation aller vom Programm erfassten Läsionen festgehalten. Diese fasste man in die Regionen Wirbelsäule, Becken, proximales Femur, proximaler Humerus, Schultergürtel, Rippenthorax



und Sternum zusammen. Die vom Algorithmus detektierten Läsionen benannte man nach ihrer Lokalisierung (z. B. proximales Femur rechts 1 und 2 bei zwei Läsionen in dieser Region). Durch anschließende Speicherung des bearbeiteten PET-Datensatzes stehen sie für eine evtl. Zweitbegutachtung zur Verfügung.

- e) Es wurde das relative metabolische Gesamtvolumen aller visuell im Referenzbogen für Skelettbefall erfassten Herde berechnet. Der Skelettbefall wurde quantifiziert, indem man einen Quotienten aus dem Volumen aller Skelettläsionen und dem Gesamtskelettvolumen (Skelettmaske) bildete.

#### 6.) Neu aufgetretene Probleme bei Testung des Algorithmus

Bei der Übertragung des anhand von elf Patienten als optimal definierten Algorithmus auf eine große Patientenkohorte (n=142) ergaben sich weitere Probleme:

- a) Spillover-Effekte durch direkt an das Skelett angrenzende, aktivitätsmehrbelegte Strukturen: Spillover bedeutet, dass ein Teil des extraskelettalen Aktivitätssignals aufgrund des Partialvolumeneffektes auf den umgebenden Hintergrund übertragen wird (Soret et al. 2007), in diesem Fall das Skelett. Dies kann z. B. bei der Harnblase, bei Lymphknotenherden mit Nähe zum Skelett, aber auch durch an Sternum bzw. Rippen thorax angrenzende Herzstrukturen oder durch aktiviertes braunes Fettgewebe um die Hals- und Brustwirbelsäule auftreten. Damit verbunden war fälschlicherweise die Detektion dieser Übertragungseffekte durch den Algorithmus als Skelettläsionen. Um dieses Problem zu lösen, wurden die Volumina dieser VOIs gesondert erfasst und nicht als Skelettbefall gewertet.
- b) Zusätzlich detektierte Läsionen: Die im Erhebungsbogen für Skelettbefall benannten Läsionen bildeten den Goldstandard zur Überprüfung der

Spezifität. Darüber hinaus wurden noch zusätzliche Läsionen detektiert. Diese rechnete man nicht zum Gesamtvolumen hinzu. Um sie aber näher zu charakterisieren, wurden ihr Volumen, die Aktivitätsbelegung und Lokalisation erfasst.

#### **5.4 Korrelation des durch den Algorithmus quantifizierten Skelettbefalls mit laborchemischen und klinischen Outcomedaten der Patienten**

Es erfolgte die Erfassung verschiedener klinischer Parameter und Laborparameter mit prognostischer Relevanz aus der Datenbank der EuroNet-PHL-C1-Studie: Hämoglobin, Gesamtprotein, Albumin, Leukozyten- und Lymphozytenzahl, Blutsenkungsgeschwindigkeit, Serumfibrinogen, Antikörperklassen IgG, IgA, IgM, Alpha-2-Globuline, Gammaglobuline, LDH, B-Symptome (Fieber, Gewichtsverlust und Nachtschweiß).

Diese Parameter wurden auf bivariate Korrelation nach Pearson mit dem quantifizierten Skelettbefall und dem Ausmaß der Knochenmarkaktivierung geprüft. Als Surrogatmarker für die Knochenmarkaktivierung diente hierbei die durchschnittliche Impulsrate des Referenzwirbelkörpers analog zu (Murata et al. 2006).

Es wurde weiterhin auf einen Zusammenhang zwischen nodalem Tumolvolumen und des anhand des vorliegenden Programmes quantifizierten Skelettbefalls auf Basis der Logarithmusskala überprüft. Die Anfertigung von Streudiagrammen diente der Veranschaulichung.

Zum Zeitpunkt der Analyse lag bei 101 von 142 Patienten eine Nachbeobachtungszeit von 36 Monaten vor. Diese Gruppe teilte man nach dem Anteil des Skelettbefalls in Bezug auf die CT-Skelettmaske etwa am Median von 0,27 Prozent. Es erfolgte die Berechnung des 36-monatigen ereignisfreien Überlebens sowie die Erstellung von Kaplan-Meier-Kurven.

## 6 Ergebnisse

### 6.1 Optimierung der Funktionsweise der vorbestehenden Programmversion (Extraktion des Skelettsystems aus dem PET-Datensatz aus der CT-Komponente)

Bei 115/142 Patienten mit Skelettbefall wurde die computergestützt erstellte Maske zur Erfassung des Gesamtskeletts (81 Prozent) durch Anwendung eines HU-Cut-Off-Wertes bei 160 erstellt. Bei 23 Patienten (16 Prozent) lag der Cut-Off zwischen 130 bis 150 HU und bei vier Patienten (3 Prozent) im Bereich von 170 bis 230 HU (Tab. 1).

Dichteschwelle (in HU)	Häufigkeit	Anteil (in %)
130	12	8,5
140	9	6,3
150	2	1,4
160	115	81,0
170	1	0,7
190	1	0,7
200	1	0,7
230	1	0,7

Tab. 1: Dichteschwelle (in HU) zur Erstellung der CT-Skelettmaske

Insgesamt hat sich die Schwelle von 160 HU bei den meisten Patienten bewährt.

Die angewandte „Knochenfüllerfunktion“ (s. g. „Hole Filling Filter“) ist in der Lage, kleinere Fehlstellen in der CT-Skelettmaske zu schließen. Größere Defekte in der CT-Skelettmaske können bisher noch nicht geschlossen werden, was bei drei Patienten die Knochenmarkhöhlen von Oberarm- und Oberschenkelknochen betraf, da deren Dichte den Schwellenwert von 160 HU deutlich unterschritt und auch eine Absenkung der Schwelle bis 130 HU nicht zur Integration der Skelettanteile in die CT-Skelettmaske führte. Das ist exemplarisch in unten stehender Abbildung 29 gezeigt, wo mit schwarzen Pfeilen markiert Anteile von proximalen Femora und Humeri, aber auch der Ossa ilia beidseits nicht in die CT-Skelettmaske integriert wurden.

Unten stehende Abbildung 30 zeigt exemplarisch einen anderen Patienten, bei dem Anteile der proximalen Femora beidseits und auch des rechten Os ischium (markiert mit schwarzen Pfeilen) nicht erfasst wurden.

Eine große Osteolyse im Schambein rechts wies bei einem weiteren Patienten eine HU-Dichte < 160 auf. Unten stehende Abbildung 31 demonstriert, dass nur die

Randbereiche dieser Läsion erfasst (gelb umrandet) und die eigentliche Osteolyse aufgrund ihrer geringen Dichte nicht in die CT-Skelettmasken integriert wurde. Derartige Probleme traten bei insgesamt drei Patienten (2 Prozent des untersuchten Kollektivs mit Skelettbefall) auf und sind somit selten.

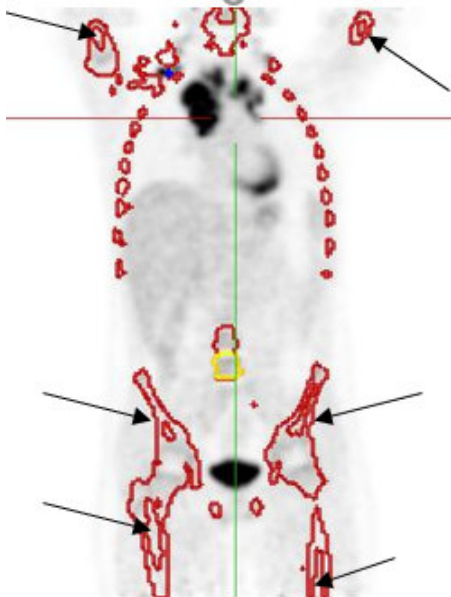


Abb. 29: Skelett mit Fehlstellen in proximalen Femora und Humeri sowie im rechten Os ilium



Abb. 30: Skelett mit Fehlstellen in Os ischium rechts und proximalen Femora

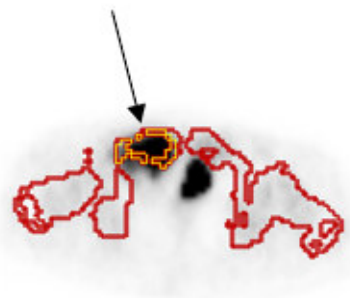


Abb. 31: Fehlstelle im rechten oberen Schambeinast durch große Osteolyse

Außerdem werden teilweise durch den „Hole Filling Filter“ die Zwischenwirbelscheiben in die CT-Skelettmasken eingeschlossen, was wiederum die Volumenbestimmung der Skelettanteile in der CT-Skelettmasken geringgradig verfälschen dürfte. Dieses Problem fällt bei rein visueller Beurteilung auf, erscheint aber aufgrund der geringen Volumina vernachlässigbar.

## **6.2 Experimentelle Ansätze zur softwaregestützten Detektion von Skelettbefall in der PET nach Programmoptimierung**

### **6.2.1 Ansatz I: Analyse des Nuklidspeicherverhaltens im extrahierten Skelett unter Normierung auf das Referenzorgan Leber**

Bei 61 untersuchten Datensätzen entsprach der Median der durchschnittlichen Aktivitätsbelegung des Leberquaders einem  $SUV_{mean}$  von 1,5 (Range 0,48 - 2,82). Die Standardabweichung der Aktivitätsbelegung des Leberquaders lag bei 0,17 (Range 0,09 - 0,38). Diese niedrige Standardabweichung spricht für die homogene Nuklidverteilung im Lebergewebe.

Der Median des Maximalwertes der Aktivitätsbelegung der Leber-VOI betrug 2,09 (Range 1,02 - 3,81), wobei kein Patient visuell eine pathologische Mehranreicherung der Leber im Sinne eines Lymphombefalls aufwies.

Es wurden nun Mittelwert mit Standardabweichung, Median mit Minimum und Maximum der Aktivitätsbelegung im Bereich der Skelettmaske zwischen den Patientengruppen mit physiologischem Skelett, Skelettbefall im initialen Staging und zum ERA (=early response assessment nach zwei Kursen OEPA-Chemotherapie) sowie mit Knochenmarkaktivierung verglichen (Tab. 2).

<b>Skelettstatus</b>	<b>Physiologisch</b>	<b>Befallen</b>	<b>Restaging</b>	<b>KM-Aktivierung</b>
Anzahl Patienten	21	10	7	13
Mittelwert $SUV_{mean}$	1,13	1,30	1,13	1,46
Median $SUV_{mean}$	1,06	1,30	1,14	1,38
SD des $SUV_{mean}$	0,37	0,24	0,30	0,40
Minimum $SUV_{mean}$	0,39	0,86	0,74	1,06
Maximum $SUV_{mean}$	2,19	1,70	1,66	2,54

Tab. 2: Kennwerte der Aktivitätsbelegungen der CT-Skelettmaske bei Patienten mit unterschiedlichem Knochenmarkstatus

Unten stehende Abbildung 32 visualisiert anhand von Boxplots die Verteilung der Aktivitätsbelegung des Skeletts zwischen diesen vier Patientengruppen.

Der gelbe Balken umfasst die Werte von der 25. bis zur 75. Perzentile, die schwarze Linie in der Mitte des Boxplot ist die Position des Medianwertes. Die „Whisker“ geben jeweils die Minimal- bzw. Maximalwerte an. Darüber hinaus sind Ausreißer außerhalb der Whisker als Punkt bzw. Sternchen markiert und mit Angabe der Patienten-ID versehen.

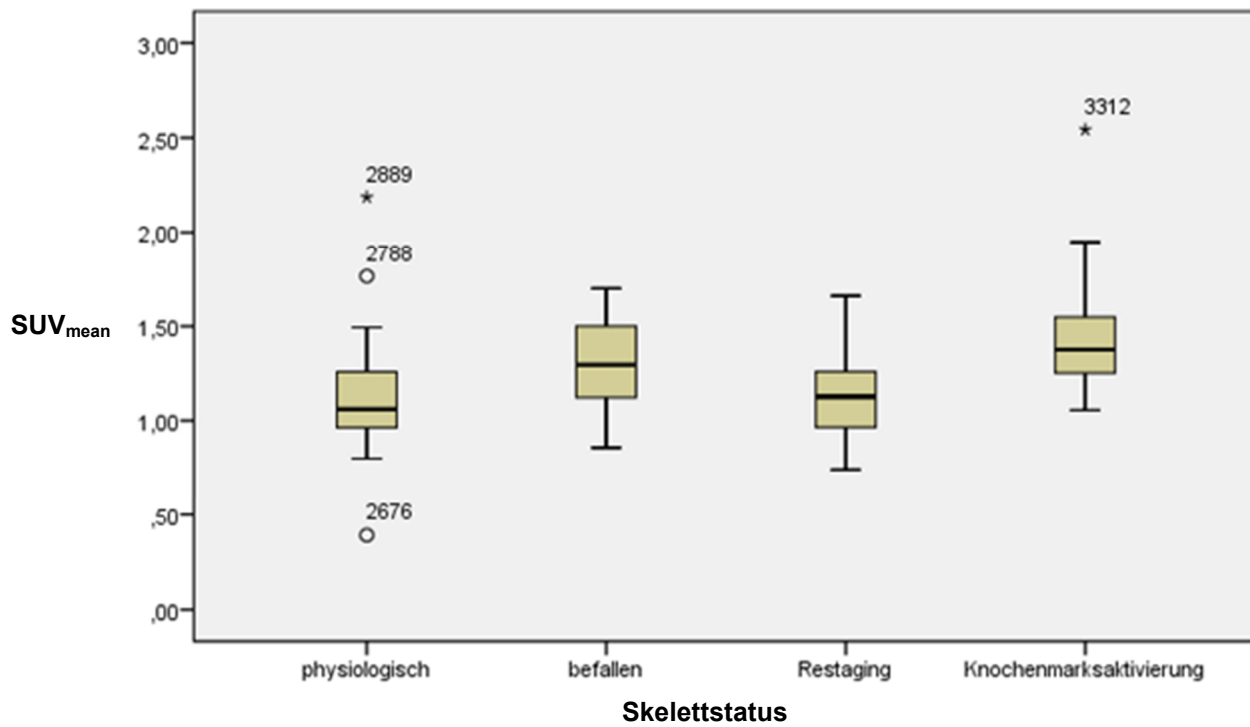


Abb. 32: Boxplots der durchschnittlichen skelettalen Aktivitätsbelegung bei unterschiedlichem Knochenmarkstatus

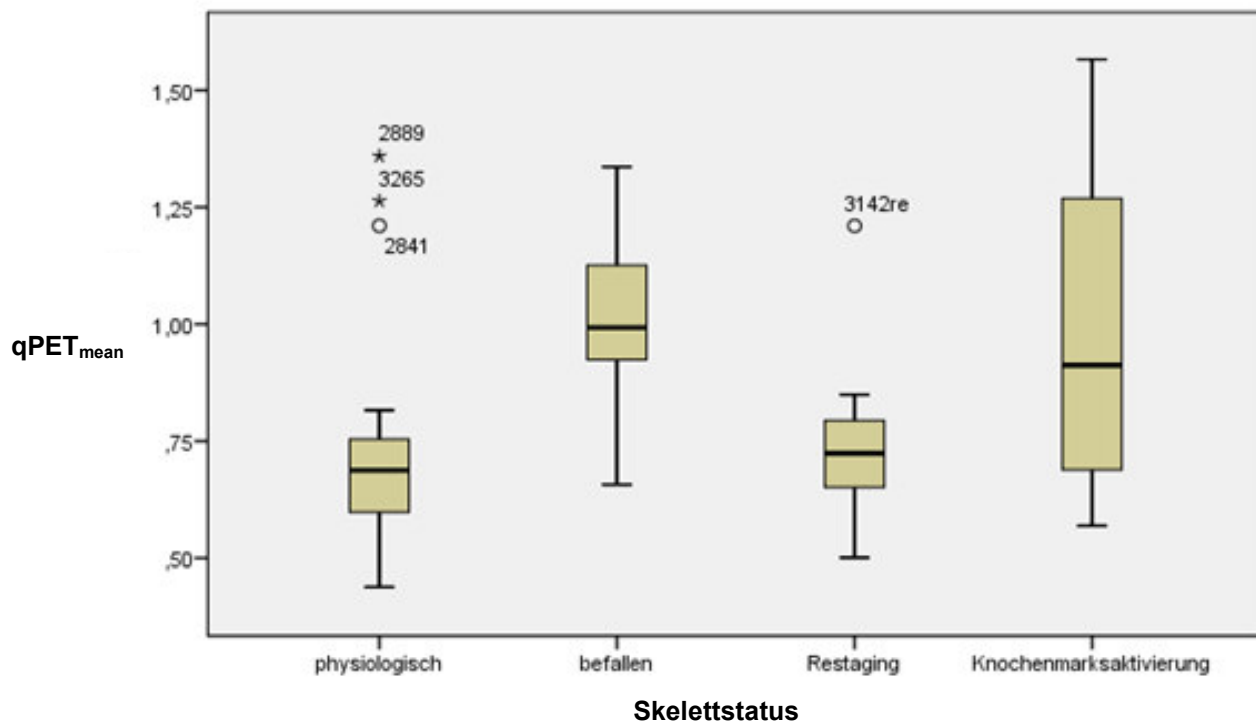


Abb. 33: Vergleich  $qPET_{mean}$ -Werte bei Patienten mit unterschiedlichem Knochenmarkstatus

Die Boxplots von physiologischem Skelettuptake und Skelettuptake im Rahmen der Frühresponsebeurteilung (ERA) erscheinen hinsichtlich ihrer Lage fast identisch, was sich auch an den nah beieinanderliegenden Medianen des  $SUV_{mean}$  von 1,06 und 1,14 zeigt.

Hingegen zeigen die Mediane der Boxplots von Patienten mit Knochenmarkbefall im Initialstaging ( $SUV = 1,30$ ) und Knochenmarkaktivierung ( $SUV = 1,38$ ) ebenso wie die Boxplots selbst eine höhere Aktivitätsbelegung.

Auffällig am Boxplot der 21 Patienten mit physiologischem Skelettuptake sind die beiden Ausreißer nach oben (\*2289,  $SUV_{mean} = 2,19$  und °2788,  $SUV_{mean} = 1,77$ ) sowie der Ausreißer nach unten (°2676,  $SUV_{mean} = 0,39$ ). Um diese Gruppe zu homogenisieren, soll nun der qPET-Wert (Quotient aus der durchschnittlichen skelettalen Aktivitätsbelegung in Bezug zur durchschnittlichen Aktivitätsbelegung des Leberquaders) gebildet werden.

Skelettstatus	Physiologisch	qPET-Wert
Mittelwert $SUV_{mean}$	1,13	0,74
Median $SUV_{mean}$	1,06	0,68
SD $SUV_{mean}$	0,37	0,25
Minimum	0,39	0,44
Maximum	2,19	1,36

Tab. 3: qPET-Werte der CT-Skelettmasken bei Patienten mit physiologischem Skelett

Durch die Quotientenbildung mit dem  $SUV_{mean}$  des korrespondierenden Leberquaders, welcher im Median 1,67 betrug, ergibt sich ein mittlerer qPET-Wert von 0,74. Die Standardabweichung wurde auf 0,25 reduziert gegenüber 0,37 vor Normierung und entspricht einer Abnahme um 32 Prozent. Der normierte Median beträgt qPET= 0,68 (Range 0,44 - 1,36). Damit verkleinert sich die Range auf 0,92 gegenüber 1,8 ohne Normierung, was einer Abnahme der Variabilität von 49 Prozent entspricht. Somit führte die Normierung auf das Referenzgewebe Leber zu einer homogenen Verteilung bei den 21 Patienten mit physiologischem Skelett.

Nun soll betrachtet werden, ob neben einer Homogenisierung auch die Diskrimination zwischen den Patientengruppen mit unterschiedlicher skelettaler Beteiligung durch Bildung des qPET-Wertes verbessert werden kann (Tab. 4).

<b>Skelettstatus</b>	<b>Physiologisch</b>	<b>Befallen</b>	<b>Restaging</b>	<b>KM-Aktivierung</b>
Anzahl Patienten	21	10	7	13
Mittelwert qPET	0,74	1,01	0,76	0,97
Median qPET	0,69	0,99	0,72	0,91
SD des SUV <sub>mean</sub>	0,25	0,17	0,23	0,33
Minimum SUV <sub>mean</sub>	0,44	0,66	0,5	0,57
Maximum SUV <sub>mean</sub>	1,36	1,34	1,21	1,57

Tab. 4: qPET-Werte der CT-Skelettmaske bei Patienten mit unterschiedlichem Knochenmarkstatus

Oben stehende Tabelle 4 zeigt, dass bei allen vier Gruppen Standardabweichung und Range verringert werden konnten, wobei auffällt, dass die Range der Gruppe mit Knochenmarkaktivierung von 0,57 bis 1,57 noch am ausgeprägtesten ist.

In dem korrespondierenden Boxplots (Abb. 33, siehe oben) umspannt die Gruppe mit diffuser Knochenmarkaktivierung hinsichtlich ihrer Range nahezu alle anderen drei Gruppen komplett. Das lässt sich am ehesten auf die unterschiedlichen Ausprägungen der Knochenmarkaktivierung zurückführen. Nach Normierung kann man rein visuell sehr gut in Abbildung 33 erkennen, dass sich die Unterscheidbarkeit zwischen Patienten mit physiologischem Skeletttuptake (Boxplot 1) bzw. sich normalisierendem Skeletttuptake im Rahmen der Frühresponsebeurteilung (max. drei Läsionen, Boxplot 3) und solchen mit mindestens fünf Läsionen (Boxplot 2) bzw. Knochenmarkaktivierung (Boxplot 4) im Vergleich zu Abbildung 32 verbessert.

Dennoch bestehen weiterhin deutliche Überlappungsbereiche zwischen den beiden letztgenannten Gruppen, die sich trotz Normierung nicht voneinander unterscheiden lassen.

Es soll des Weiteren der Rückgang der Aktivitätsbelegung zwischen Initialstaging und der Frühresponsebeurteilung nach zwei Zyklen Chemotherapie näher betrachtet werden. Von besonderem Interesse erscheint die Unterscheidbarkeit der Aktivitätsbelegung bei geringem Restbefall gegenüber Patienten ohne Skelettbefall.

Von den sieben ausgewählten Patienten mit initialem Skelettbefall waren bei drei Patienten noch jeweils zwei Herde gegenüber vormals 13, 19 bzw. 26 Herden, bei einem Patienten noch drei Herde gegenüber 27 Herden und bei drei Patienten keine Läsionen bei initial 18, 19 bzw. 34 Herden im Skelett nachweisbar.



Unten stehende Tabelle 5 zeigt den deutlichen Rückgang des Medianwertes des  $SUV_{mean}$  im Skelett von vormals 1,24 im Initialstaging auf 1,14 zur Frühresponsebeurteilung (ERA nach zwei Zyklen OEPA-Chemotherapie).

Patienten mit Skelettbefall	Initialstaging	ERA
Anzahl Patienten	7	7
Median $SUV_{mean}$	1,24	1,14
Median Leberquader $SUV_{mean}$	1,23	1,38
Median qPET-Wert	1,08	0,72

Tab. 5: SUV- und qPET-Werte der CT-Skelettmassage bei 7 Patienten mit Skelettbefall im Initialstaging und ERA-PET

Bildet man erneut den qPET-Wert als Quotient aus durchschnittlichem Skelettuptake und Aktivitätsbelegung des Leberquaders, wird dieser Unterschied noch deutlicher. Von initial 1,08 im Initialstaging reduziert sich der qPET-Wert auf 0,72 im Restaging, was einer Reduktion um ca. ein Drittel entspricht und sich auch in den Boxplots unten stehender Abbildung 34 deutlich zeigt. Hierzu trägt ein Anstieg des Leber-SUV bei. Der Ausreißer 3142re ergibt sich, weil am ehesten durch eine geringe Patientenbewegung zwischen CT und PET-Akquisition eine Überlagerung des Schädelknochens (aus der CT-Komponente) mit dem Gehirn (aus der PET-Komponente) erfolgte (siehe unten, Abbildung 44), wodurch die Aktivitätsbelegung deutlich erhöht wurde.

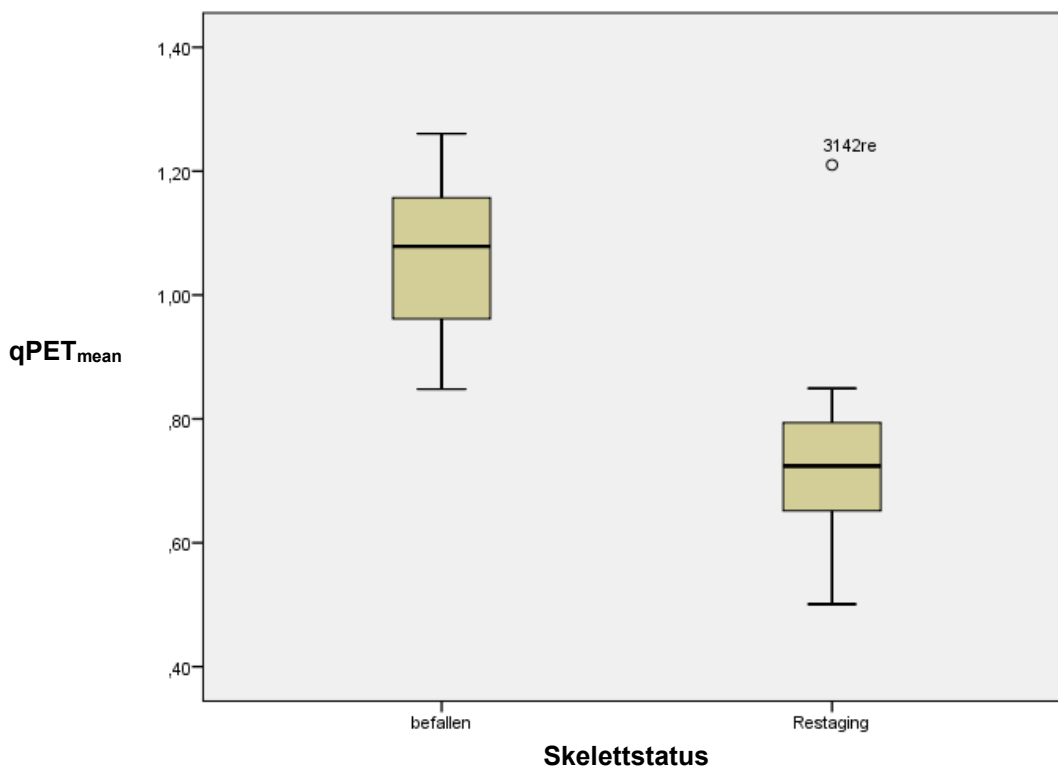


Abb. 34: Normierter skelettaler Uptake der sieben Patienten im Initialstaging und ERA

Vergleicht man stichprobenhaft die drei Patienten im ERA ohne residuellen Skelettbefall mit den Patienten mit zwei bzw. drei residuellen Herden (Tab. 6), so unterscheiden sich die Mediane der qPET-Werte hingegen kaum (0,72 vs. 0,73). Das spricht dafür, dass ein wenig ausgedehnter Skelettbefall in der ERA-PET nicht von Patienten ohne Skelettbefall bei Analyse der gesamten CT-Skelettmaske unterscheidbar ist, wobei die Restbefunde im ERA nur noch geringe SUV aufweisen.

Restaging	Kein residueller Skelettbefall	Geringer residueller Skelettbefall
Anzahl Patienten	3	4
Median Skelett-SUV <sub>mean</sub>	1,00	1,23
Median Leber-SUV <sub>mean</sub>	1,38	1,57
qPET-Wert	0,72	0,73

Tab. 6: qPET-Werte der CT-Skelettmaske bei Patienten im ERA-PET ohne und mit geringem residuellem Befall

Da sich die Aktivitätsbelegung des gesamten Skeletts trotz Normierung auf das Referenzorgan Leber bei Knochenmarkaktivierung (Mittelwert qPET = 1,01; SD +/- 0,17) nicht wesentlich von der Gruppe mit Skelettbefall (Mittelwert qPET = 0,97; SD +/- 0,33) unterscheidet, wurden weitere Methoden zur Erkennung von Skelettbefall untersucht.

### 6.2.2 Ansatz II: Histogrammanalyse

Die Histogrammdarstellungen, die die normierte Aktivitätsbelegung auf der Abszisse gegenüber dem Anteil der jeweiligen Voxel mit dieser Aktivitätsbelegung am Gesamtskelett auf der Ordinate abbilden, wurden zunächst visuell miteinander verglichen. Die beiden unten stehenden Abbildungen zeigen Histogramme zweier Patienten mit physiologischem Nuklidbesatz des Skeletts (Abb. 35, 36). Das Histogramm der Abbildung 35 zeigt hierbei annähernd eine Normalverteilung mit einem Peak des qPET-Wertes bei 0,45.

Nur sehr wenige Skelettvoxel überschreiten den Durchschnittswert der Leber und damit einen qPET-Wert von 1. Das Histogramm in Abbildung 36 dagegen ist mehr nach rechts verbreitert, der Peak des qPET-Wertes liegt zwischen 0,5 bis 0,6.

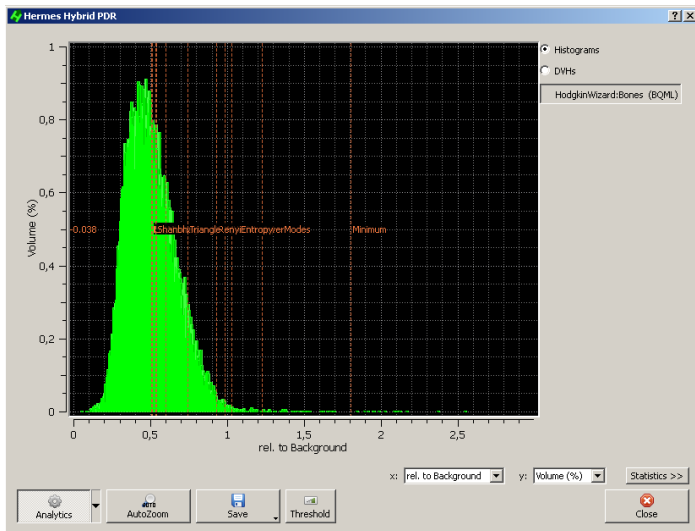


Abb. 35: Beispielhistogramm 1 ohne Skelettbefall oder Knochenmarkaktivierung

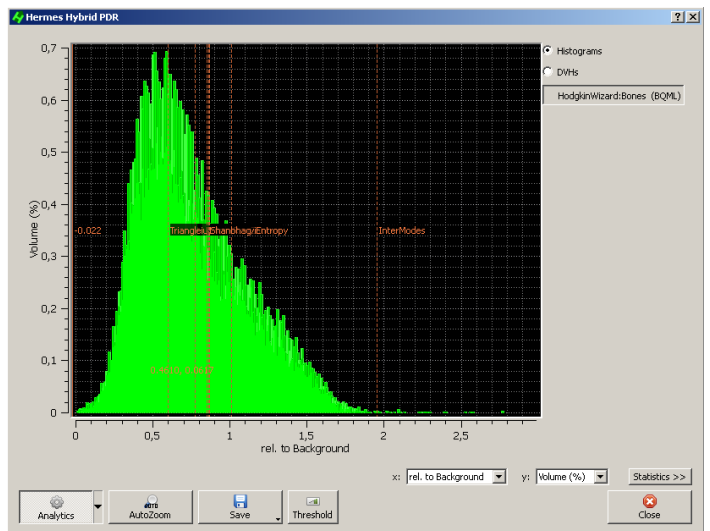


Abb. 36: Beispielhistogramm 2 ohne Skelettbefall oder Knochenmarkaktivierung

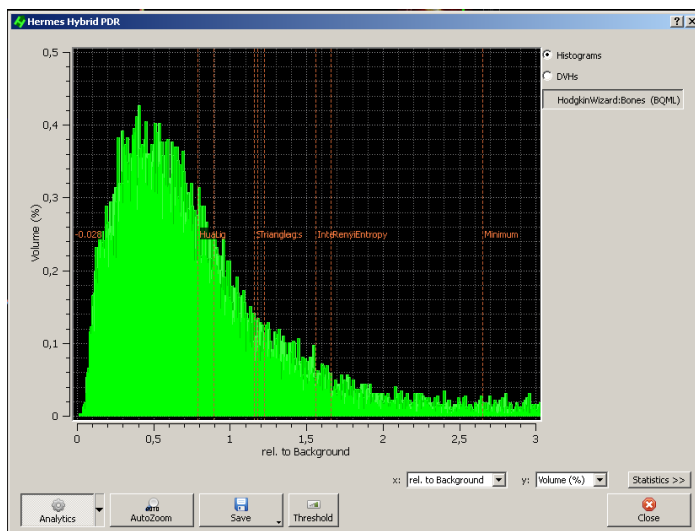


Abb. 37: Beispielhistogramm 1 mit ausgeprägtem Skelettbefall

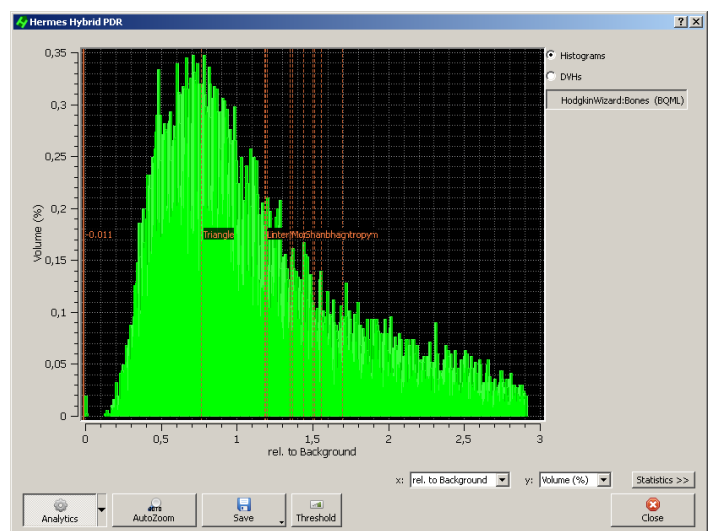


Abb. 38: Beispielhistogramm 2 mit ausgeprägtem Skelettbefall

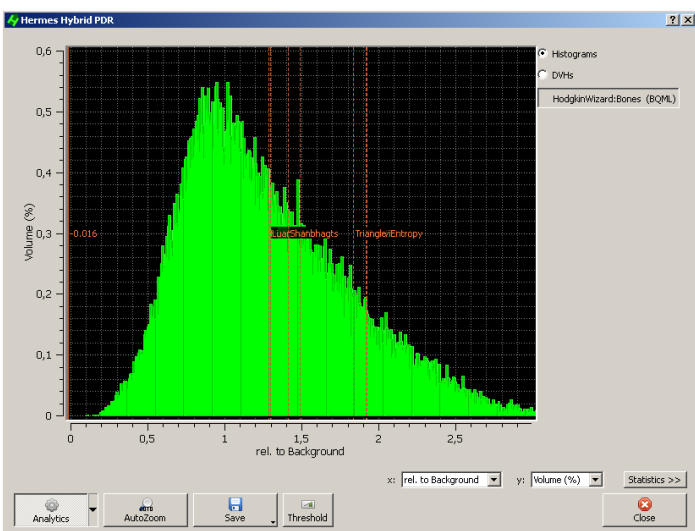


Abb. 39: Beispielhistogramm 1 mit Knochenmarkaktivierung

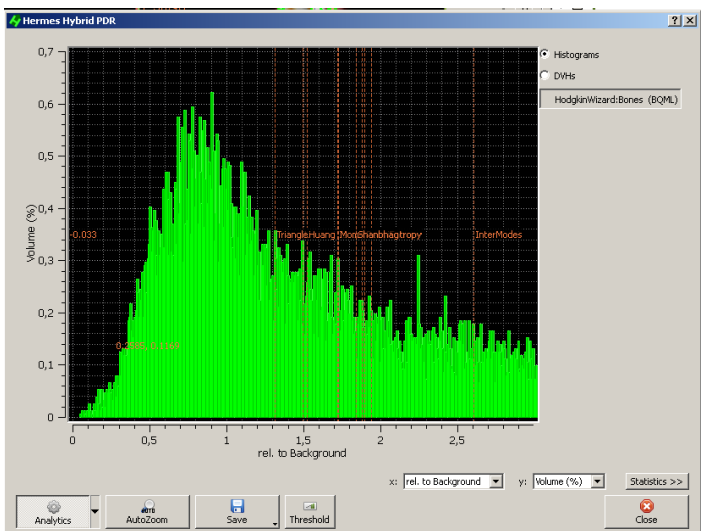


Abb. 40: Beispielhistogramm 2 mit Knochenmarkaktivierung

Auch wenn sich die Konfigurationen dieser Histogramme unterscheiden, so sind sie doch von solchen mit Skelettbefall oder Knochenmarkaktivierung (Abb. 37 bis 40) abgrenzbar.

Das illustrieren oben stehende Abbildungen 37 und 38, welche jeweils Patienten mit einem ausgeprägtem Skelettbefall zeigen. Bei Abbildung 37 fällt auf, dass der Peak des qPET-Wertes wie bei den vorher gezeigten Diagrammen (Abb. 35, 36) bei etwa 0,5 liegt. Allerdings fällt die Dichtekurve dann selbst jenseits der doppelten Leberaktivitätsbelegung nicht weiter ab. Zudem sind Voxel mit einer bis zu dreifachen Leberaktivitätsbelegung vorhanden. Abbildung 38 zeigt hingegen insgesamt eine nach rechts verbreiterte Darstellung mit einem Peak bei 0,8. Nach dem Peak fällt die Kurve nur langsam ab. Auch wenn die Konfiguration dieser Histogramme nicht einheitlich ist, so weisen doch beide als Unterscheidungsmerkmal zu den Histogrammen bei physiologischem Skelettuptake eine vergleichsweise ausgeprägte Rechtsverbreiterung auf; außerdem gibt es zahlreiche Voxel mit mehr als der doppelten Aktivitätsbelegung der Leber.

Exemplarisch sind noch zwei Histogramme von Patienten mit diffuser Knochenmarkaktivierung, jedoch ohne Skelettbefall demonstriert (oben stehende Abb. 39, 40). In Abbildung 39 liegt der Peak des qPET-Wertes bei etwa 0,9. Analog zu den Abbildungen 37 und 38 fällt auch hier eine deutliche Rechtsverbreiterung mit nur langsamem Kurvenabfall nach dem Peak und einer größeren Voxelzahl mit mehr als der doppelten Leberaktivitätsbelegung auf. In Abbildung 40 liegt der Peak bei ca. 0,8. Es fällt aber auf, dass der Kurvenabfall nach dem Peak eher dem Übergang in ein Plateau ähnelt, das wiederum zahlreiche Voxel mit mehr als der doppelten Leberaktivitätsbelegung beinhaltet.

Nach visueller Analyse der Histogramme kann als Schlussfolgerung gezogen werden, dass eine Differenzierung zwischen Skelettbefall und Knochenmarkaktivierung aufgrund breiter Überlappungsbereiche nicht gelingt.

Als weitere Möglichkeit der Generierung von Unterscheidungskriterien wurden verschiedenste Histogramm-Kenngrößen (s. g. „Analytics“) durch eine im Programm implementierte Funktion berechnet und im Histogramm angezeigt. Wie aus Tabelle 7 zu entnehmen ist, sind Analytics ohne entsprechende Normierungen nicht zu verwerten. Hier sind die Kilocounts pro Sekunde für elf berechnete Analytics von zehn

Patienten ohne Skelettbefall (ID) aufgetragen. Die erfassten Analytics zeigen eine ausgeprägte Variabilität ohne jegliche Systematik.

ID	Hua	Inter	Min	Isod	Li	Max	Mom	Yen	Ren	Sha	Tria
3200	6,5	18,0	32,0	9,1	7,0	10,4	9,1	10,4	10,4	22,0	6,0
3290	1,8	19,0	30,0	2,4	2,0	6,3	4,4	6,3	6,3	15	3,5
3161	3,0	12,0	22,0	7,0	5,0	9,0	6,0	8,0	8,0	14,0	5,0
3208	3,9	32,0	48,0	4,2	4,0	11,5	8,0	12,2	11,5	18	6,0
3202	3,0	9,0	13,5	3,9	3,5	7,0	5,0	7,0	6,8	7,0	5,5
2841	3,1	3,4	5,3	3,8	2,9	4,2	4,1	3,5	3,6	8,6	2,5
2847	3,4	22,0	41,0	9,5	6,2	11,4	9,6	8,6	8,7	26,0	6,0
3252	3,6	6,8	10,6	4,0	3,6	4,9	4,0	4,8	4,8	4,1	4,2
2810	1,9	6,0	10,0	2,0	1,9	4,1	2,4	4,2	3,5	2,0	2,0
2815	1,5	5,4	8,9	1,8	1,7	5,1	2,2	5,8	3,8	1,8	2,7

Tab. 7: Position der Analytics mit Kilocounts als Abszisse im Histogramm bei Patienten mit physiologischem Skelett

Vergleicht man nun die Position der Analytics anhand des qPET-Wertes als Quotient aus Voxelaktivitätsbelegung und Leberaktivitätsbelegung, so ergibt sich ein etwas homogeneres Bild (Tab. 8). Aber auch nach diesem Normierungsschritt ist keiner der aufgeführten Analytics in der Lage, Grenz- oder Schwellenwerte zu beschreiben.

ID	Hua	Inter	Min	Isod	Li	Max	Mom	Yen	Ren	Sha	Tria
3200	0,35	1,0	1,8	0,5	0,4	0,6	0,5	0,6	0,6	1,25	0,35
3290	0,25	2,9	4,6	0,35	0,3	0,95	0,67	0,95	0,95	2,3	0,55
3161	0,2	2,4	4,4	1,4	1,0	1,8	1,2	1,6	1,6	2,8	1,0
3208	0,75	6,1	9,2	0,8	0,75	2,2	1,5	2,3	2,2	3,45	1,1
3202	0,6	1,9	3,0	0,85	0,8	1,55	1,1	1,55	1,5	1,55	1,2
2841	0,6	0,65	1,0	0,75	0,55	0,55	0,8	0,7	0,7	1,7	0,5
2847	0,25	1,8	3,35	0,8	0,5	0,9	0,8	0,7	0,7	2,1	0,5
3252	0,5	0,95	1,5	0,55	0,5	0,7	0,55	0,7	0,7	0,6	0,6
2810	0,55	1,7	2,2	0,6	0,55	1,2	0,7	1,2	1,0	0,6	0,6
2815	0,6	2,3	3,7	0,75	0,7	2,1	0,9	2,4	1,6	0,7	1,1

Tab. 8: Position der Analytics mit qPET-Wert als Abszisse im Histogramm bei Patienten mit physiologischem Skelett

Die vorgestellten Ergebnisse zur Histogrammanalyse sind wenig vielversprechend, sodass bei Vorhandensein alternativer Lösungsansätze die Fortentwicklung dieses methodischen Ansatzes im Rahmen dieser Dissertationsarbeit nicht weiterverfolgt wurde.

Der Zeitaufwand für die Erstellung der CT-Skelettmaske, Histogrammgenerierung, einheitliche Formatierung und Speicherung sowie manuelle Erfassung der histogrammbasierten Schwellenwerte (Analytics) betrug etwa 40 Minuten pro Patient.

### **6.2.3 Ansatz III: Aktivitätsanalyse ausgewählter Skelettregionen mit und ohne Lymphommanifestationen**

#### **a) Patienten im Initialstaging mit Skelettbefall**

Um betrachten zu können, wie Skelettläsionen die Aktivitätsbelegung der verschiedenen Skelettregionen quantitativ beeinflussen, wurden zunächst elf Patienten im Initialstaging mit insgesamt 222 Skelettläsionen betrachtet. Getrennt nach Skelettregionen sind die Durchschnitte der  $qPET_{mean}$ - und  $qPET_{max}$ -Werte der befallenen Skelettabschnitte zusammengestellt (Tab. 9).

Wie bereits im Methodikteil erläutert, befinden sich besonders viele Läsionen in Wirbelsäule (81 von 222 Läsionen) und Becken (101 von 222 Läsionen).

Dies entspricht insgesamt einem Anteil von 88,3 Prozent. Tabelle 9 zeigt passend dazu eine Erhöhung des Durchschnitts von  $qPET_{mean}$  im Becken über 1,41 und in der Wirbelsäule über 1,29.

Auf der anderen Seite beherbergen obere und untere Extremität sowie das Thoraxskelett nur 11,7 Prozent aller Skelettläsionen. Hier finden sich deutlich niedrigere  $qPET$ -Werte von 0,7 bis 0,9.

Auffällig sind im Gegensatz dazu die deutlich erhöhten Durchschnittswerte des Schädels von  $qPET_{mean} = 1,16$  und der  $qPET_{max} = 7,02$ , obwohl kein Schädelbefall vorlag.

Region	qPET <sub>mean</sub>	qPET <sub>max</sub>
Wirbelsäule	1,29	7,61
Rippenthorax	0,70	3,96
Becken	1,41	9,61
Humeri mit Skapula	0,71	4,89
Femora	0,90	5,05
Schädel	1,16	7,02

Tab. 9: Durchschnitt der qPET-Werte, getrennt nach Skelettregionen bei Patienten mit Skelettbefall im Initialstaging

Abbildung 41 stellt nun die Boxplots der qPET<sub>mean</sub> der einzelnen Skelettregionen dar. Hier zeigt sich, korrespondierend zu Tabelle 9, dass die Boxplots von Becken und Wirbelsäule deutlich höhere Werte als in oberer und unterer Extremität bzw. Thoraxskelett annehmen.

Auch hier fällt auf, dass der Schädel einen Boxplot in ähnlicher Position wie die Boxplots von Becken und Wirbelsäule zeigt, obwohl kein Schädelknochen bei den elf untersuchten Patienten mit Skelettbefall eine Läsion in der Referenzbegutachtung aufwies.

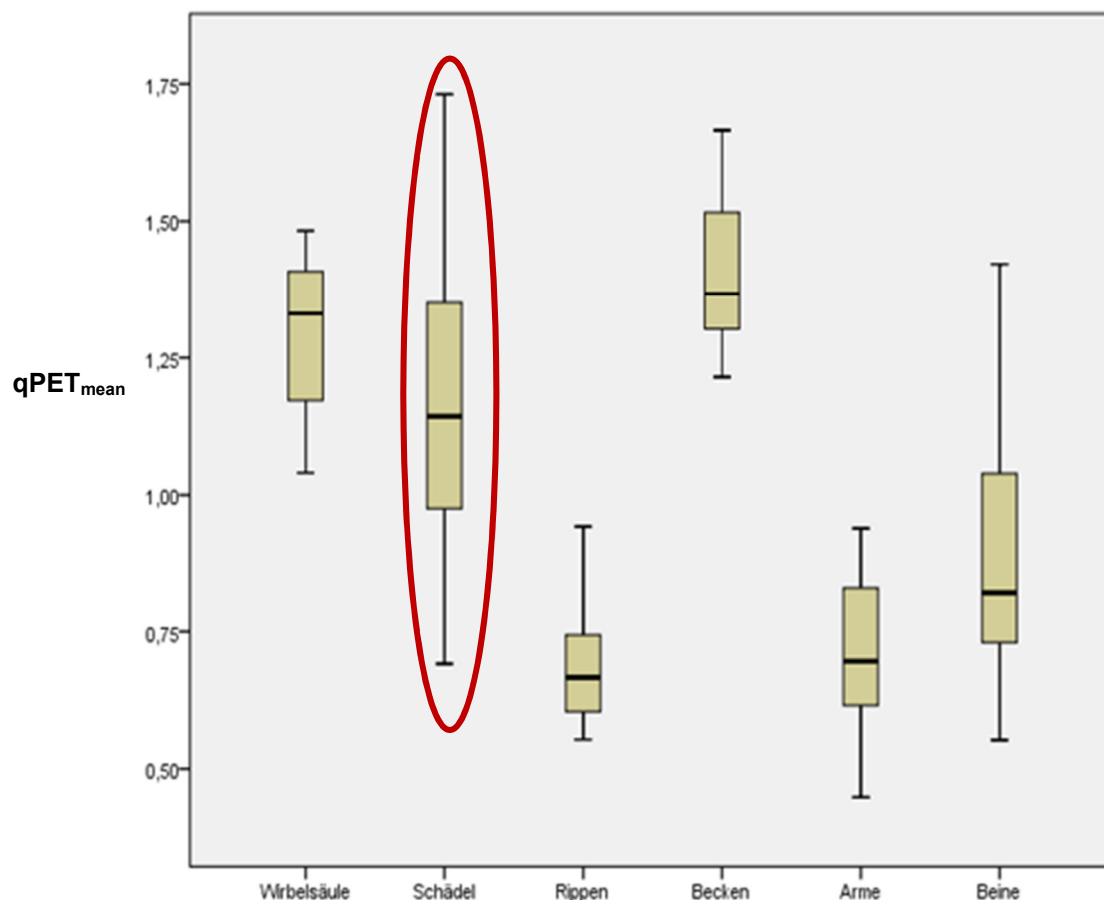


Abb. 41: Boxplots der qPET<sub>mean</sub>-Werte bei Skelettbefall im Initialstaging, getrennt nach Skelettregionen

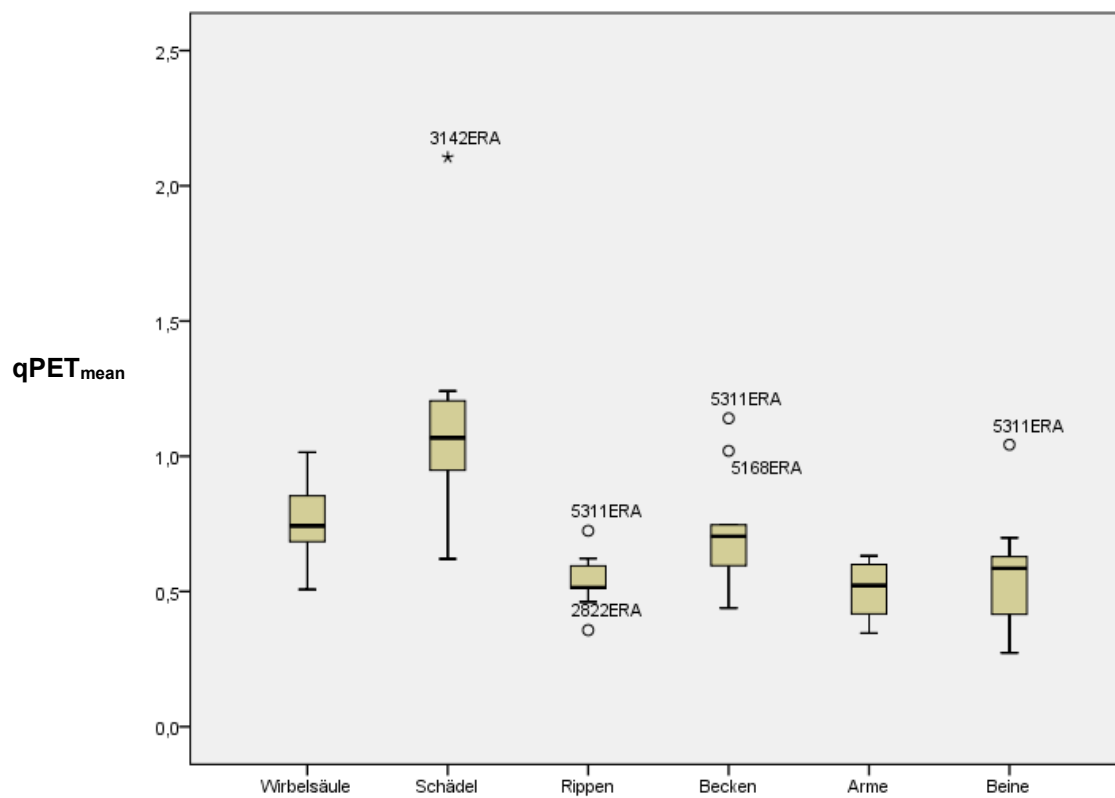


Abb. 42: Boxplots der  $qPET_{mean}$ -Werte im ERA-PET, getrennt nach Skelettregionen

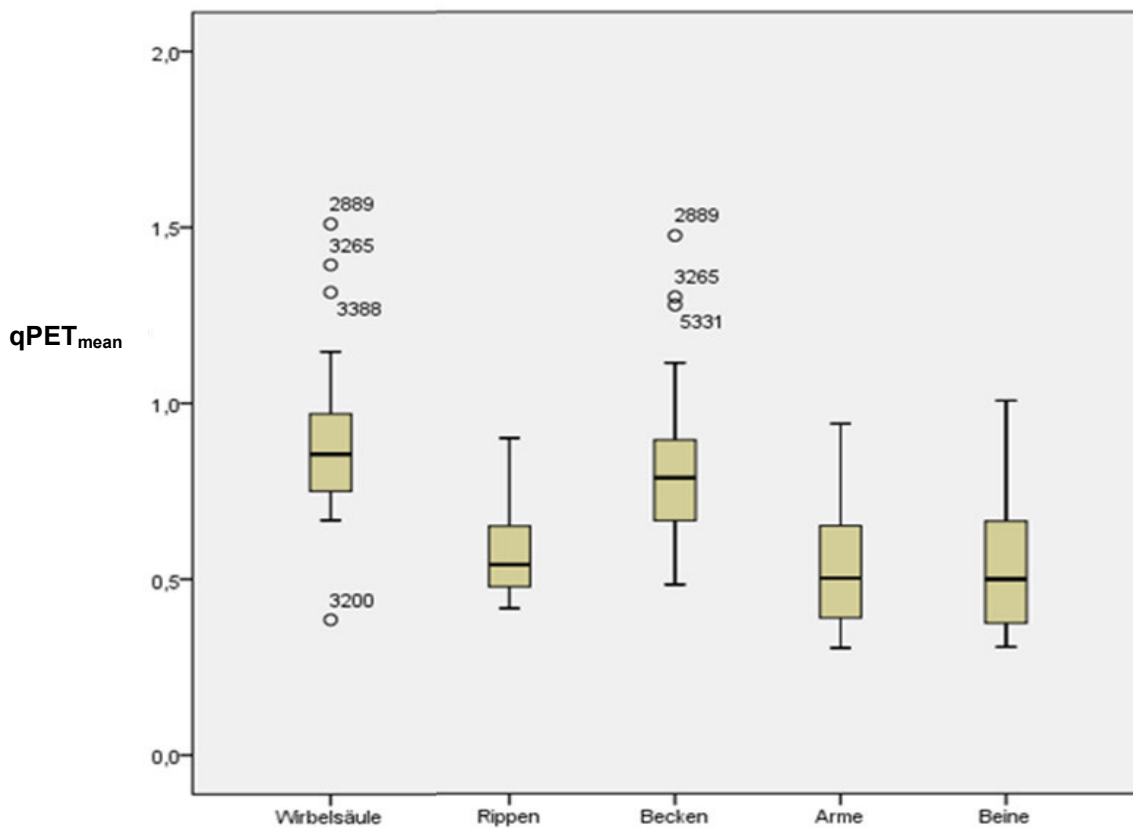


Abb. 43: Boxplots der  $qPET_{mean}$ -Werte ohne Skelettbefall im Initialstaging, getrennt nach Skelettregionen



b) Patienten zur Frühresponsekontrolle nach zwei Kursen OEPA-Chemotherapie

Bei zehn der oben genannten elf Patienten wurde die zusätzlich verfügbare Frühresponsekontrolle (ERA nach zwei Zyklen OEPA) dahingehend betrachtet, ob sich ein Ansprechen des Skelettbefalls auf die Chemotherapie auch in einer Verringerung der qPET-Werte der jeweiligen Skelettregion zeigt. Beim elften Patienten lag die Frühresponsekontrolle zum Zeitpunkt der Auswertung nicht vor.

Es ergaben sich die in Tabelle 10 aufgeführten qPET<sub>mean</sub>- und qPET<sub>max</sub>-Werte der jeweiligen Regionen. Des Weiteren wurde die prozentuale Veränderung der mittleren qPET-Werte zwischen initialem Staging und ERA ermittelt. Es zeigt sich, dass in den jeweiligen Regionen die Durchschnittswerte von qPET<sub>mean</sub> um 22,9 Prozent bis 48,3 Prozent abnehmen und von qPET<sub>max</sub> um 39,9 Prozent bis 64,5 Prozent.

Die deutlichste Reduktion des qPET<sub>mean</sub> um 48,3 Prozent findet sich im Bereich des Beckens, der Region mit den initial meisten Skelettläsionen.

Deutliche Abweichungen lassen sich erneut im Bereich des Schädels feststellen. Hier weist der Durchschnitt des qPET<sub>mean</sub>-Wertes mit nur 3,6 Prozent keine wesentliche Reduktion auf. Der Durchschnittswert des qPET<sub>max</sub>-Wertes ist gar um 39,3 Prozent im Vergleich zum initialen Staging angestiegen, was nicht plausibel ist.

Region	qPET <sub>mean</sub>	Reduktion qPET <sub>mean</sub> zum Initialstaging um	qPET <sub>max</sub>	Reduktion qPET <sub>max</sub> zum Initialstaging um
Wirbelsäule	0,77	31,3 %	2,70	64,5 %
Thorax mit Rippen	0,54	22,9 %	2,38	39,9 %
Becken	0,73	48,3 %	3,50	63,6 %
Humeri mit Skapulae	0,51	28,2 %	2,04	58,2 %
Femora	0,56	37,8 %	2,15	57,4 %
Schädel	1,13	3,6 %	9,78	-39,3 %

Tab. 10: qPET-Werte, getrennt nach Skelettregionen bei Patienten im ERA und Veränderung gegenüber dem Initialstaging

Es sollte des Weiteren überprüft werden, ob sich im ERA die beiden Gruppen (geringe Residuen bei initialem Skelettbefall vs. keine Residuen bei initialem Skelettbefall) anhand des Uptakes in den jeweiligen Skelettregionen unterscheiden lassen. Bei sieben der zehn Patienten war das Skelett in der Frühresponsebeurteilung als lymphomfrei bewertet worden. Bei drei Patienten waren noch drei Läsionen in der Wirbelsäule, vier Läsionen im Becken, eine Läsion im Femur, aber keine in Humerus oder Thoraxskelett vorhanden.

Vergleicht man die Reduktion des  $qPET_{\text{mean}}$  gegenüber dem Initialstaging der drei Patienten mit einzelnen residuellen Läsionen in Wirbelsäule, Becken und Femora (Tab. 11) mit der Gesamtgruppe aller zehn ERA-Patienten (Tab. 10), so fällt auf, dass diese in den jeweiligen Regionen keinen wesentlichen Unterschied zeigen (Wirbelsäule 43,1 Prozent gegenüber 41,8 Prozent, Becken 53,5 Prozent gegenüber 48,3 Prozent, Femora 36,3 Prozent gegenüber 33,8 Prozent).

Region	$qPET_{\text{mean}}$ beim ERA	$qPET_{\text{mean}}$ beim Initialstaging	Reduktion vom ERA zum Initialstaging um
Wirbelsäule	0,68 (0,51-0,85)	1,45 (1,36-1,48)	43,1 %
Becken	0,66 (0,44-0,69)	1,42 (1,29-1,67)	53,5 %
Femora	0,47 (0,27-0,6)	0,75 (0,55-1,13)	36,3 %

Tab. 11: Median und Range der  $qPET_{\text{mean}}$ -Werte, getrennt nach Skelettregionen bei Patienten mit residuellem Befall im ERA und Veränderung gegenüber dem Initialstaging

Auch die Maximalwerte der jeweiligen Skelettregion zeigen eine deutliche Reduktion um mindestens 64,7 Prozent (Tab. 12) bei residuellem Skelettbefall gegenüber mindestens 57,4 Prozent (Tab. 10) bei allen Patienten.

Region	$qPET_{\text{max}}$ beim ERA	$qPET_{\text{max}}$ beim Initialstaging	Reduktion vom ERA zum Initialstaging um
Wirbelsäule	1,68 (1,44-1,87)	5,63 (5,04-12,04)	70,2 %
Becken	2,01 (1,80-2,25)	7,93 (3,69-14,08)	74,7 %
Femora	1,17 (0,78-1,40)	3,31 (3,13-6,38)	64,7 %

Tab. 12: Median und Range der  $qPET_{\text{max}}$ -Werte, getrennt nach Skelettregionen bei Patienten mit residuellem Befall im ERA und Veränderung gegenüber der Initialstaging

Die Reduktion der Aktivitätsbelegung ist somit bei geringem residuellen Befall in der jeweiligen Skelettregion sogar etwas höher als bei vollständigem Ansprechen der Skelettmanifestationen auf die Chemotherapie in der hier untersuchten kleinen Stichprobe.

Ein geringer residueller Befall lässt sich somit nicht von einer vollständigen Remission über den Vergleich der Aktivitätsbelegung der jeweiligen Skelettregion unterscheiden.

Betrachtet man die Boxplots aller zehn Patienten und verfügbarem ERA (Abb. 42, siehe oben) gegenüber dem Initialstaging (Abb. 41, siehe oben), so lässt sich bereits visuell die deutliche Reduktion der  $qPET_{\text{mean}}$ -Werte zwischen den beiden Untersuchungsanlässen erkennen. Die Boxplots aller Regionen mit Ausnahme des Schädels weisen einen  $qPET_{\text{mean}}$ -Wert  $< 1$  auf.

Es soll nun die Ursache für die deutlich erhöhten  $qPET$ -Werte des Schädels bei fehlendem Skelettbefall betrachtet werden.

Der Mittelwert von  $qPET_{\text{mean}}$  war im Vergleich zum Initialstaging nahezu unverändert und der Mittelwert von  $qPET_{\text{max}}$  hatte sogar um 39 Prozent zugenommen (Tab. 10), obwohl nach visueller Referenzbegutachtung ein physiologischer Tracerbesatz der Schädelkalotte ohne skelettale Läsionen vorlag: Hierzu ist anzumerken, dass das Gehirn in der FDG-PET durch seinen physiologisch erhöhten Glukoseumsatz auch eine entsprechend starke Anreicherung des Radiotracers ( $^{18}\text{F}$ -FDG) aufweist. Durch Spillover-Effekte kommt es bei der Übertragung der CT-Skelettmaske auf den PET-Datensatz zur Projektion von Hirnaktivität auf die Schädelkalotte und somit zu falsch hohen Aktivitätswerten. Auch führen schon diskrete Kopfbewegungen des Patienten zur fehlerhaften Koregistrierung der Datensätze des Schädels aus PET und CT.

Abbildung 44 zeigt beispielhaft, wie nach Übertragung der CT-Skelettmaske auf den PET-Datensatz Hirnaktivität auf die Schädelkalotte im Bereich der Frontobasis und des Okziputs projiziert wird.

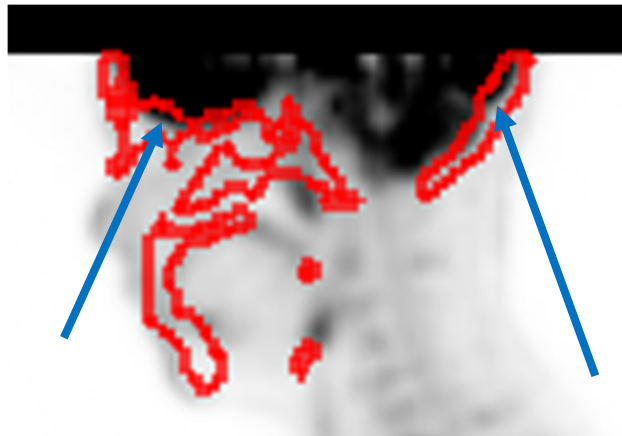


Abb. 44: Spillover der Hirnaktivität auf die Schädelkalotte

Die Schädelkalotte ist somit nicht durch Übertragung einer aus dem CT-Datensatz generierten Skelettmaske auf den PET-Datensatz be- und auswertbar. Um fehlerhafte Messwerte zu vermeiden, wurde der Schädel in weiterführenden Auswertungen nicht mehr in die endgültige CT-Skelettmaske integriert. Dieses Vorgehen ist legitim, da der Schädel häufig überhaupt nicht vollständig abgebildet ist und zudem ein Befall der Schädelknochen beim Hodgkin-Lymphom äußerst selten vorkommt.

#### c) Patienten im Initialstaging ohne Skelettbefall

Bei 18 Patienten ohne Skelettbefall sollte untersucht werden, ob regionenspezifische Schwellenwerte zur Unterscheidung zwischen Skelettbefall und physiologischer Tracerbelegung des Skeletts nötig sind. In Tabelle 13 sind die unnormierten Durchschnittswerte von  $SUV_{mean}$  und  $SUV_{max}$  der nach Ausschluss des Schädels verbliebenen fünf Regionen gegenübergestellt worden. Hierbei fällt eine um etwa 50 Prozent höhere, mittlere Aktivitätsbelegung von Wirbelsäule und Becken im Vergleich zu Thoraxskelett und Extremitäten auf.

Region	$SUV_{mean}$	$SUV_{max}$
Wirbelsäule	1,59	4,15
Rippenthorax	0,99	4,12
Becken	1,43	5,54
Humeri mit Skapulae	0,89	3,96
Femora	0,91	3,34

Tab. 13: Mittlere und maximale SUV-Werte, getrennt nach Skelettregionen bei Patienten ohne Skelettbefall

Die anschließend durchgeführte Normierung auf den durchschnittlichen Uptake des Leberquaders ergab die in Tabelle 14 dargestellten qPET-Werte, jeweils getrennt nach unterschiedlichen Skelettregionen. Der Unterschied zwischen den beiden oben beschriebenen Gruppen von Skelettregionen (Becken, Wirbelsäule vs. obere/untere Extremität, Thorax) bleibt aber im Wesentlichen bestehen. Die Aktivitätsbelegung des Schädels ist hierbei durch Spillover-Effekte des Gehirns falsch erhöht, was im vorhergehenden Abschnitt b) herausgearbeitet wurde. Soll ein einheitlicher Schwellenwert für alle Skelettregionen verwendet werden, so müsste dieser idealerweise so gewählt werden, dass pathologische Mehrbelegungen der Skelettabschnitte mit geringerer Impulsdichte wie Thoraxskelett und Extremitäten sicher erkannt werden, ohne dass die physiologisch höheren Aktivitätsbelegungen der Skelettabschnitte Becken und Wirbelsäule als falsch-positive Läsionen des Hodgkin-Lymphoms fehlgedeutet werden.

Region	qPET <sub>mean</sub>	qPET <sub>max</sub>
Wirbelsäule	0,91	2,41
Rippenthorax	0,56	2,48
Becken	0,83	3,36
Humeri mit Scapulae	0,53	2,45
Femora	0,53	1,78

Tab. 14: Mittlere und maximale qPET-Werte, getrennt nach Skelettregionen bei Patienten ohne Skelettbefall

Abbildung 43 (siehe oben) visualisiert die Boxplots mit den qPET<sub>mean</sub>-Werten der 18 Patienten ohne Skelettbefall. Vergleicht man diese mit den Boxplots der elf Patienten mit Skelettbefall (Abb. 41), so fällt auf, dass diese sich deutlich unterscheiden. So liegen hier die 75%-Perzentilen sämtlich unter einem qPET-Wert von 1, ähnlich wie in den Boxplots der Patienten in der Frühresponsebeurteilung (ERA nach 2x OEPA, siehe Abb. 41).

Tabelle 15 stellt die 75%-Perzentilen der qPET<sub>mean</sub>-Werte der physiologischen Skelettregionen den 25%-Perzentilen der qPET<sub>mean</sub>-Werte von lymphombefallenen Skelettregionen gegenüber, wie sie die Abbildungen 41 und 43 visualisieren. Hierbei zeigt sich deutlich, dass sich insbesondere die qPET<sub>mean</sub>-Werte der befallenen Regionen mit vielen Läsionen wie z.B. Becken und Wirbelsäule (siehe Methodikteil

Kapitel 5.2.3) besser als die Regionen mit einer geringen Anzahl von Läsionen (Rippenthorax, Humeri und Femora) unterscheiden lassen.

Region	75%-Perzentile $qPET_{mean}$ Physiologisches Skelett	25%-Perzentile $qPET_{mean}$ Skelettbefall
Wirbelsäule	0,97	1,18
Rippenthorax	0,65	0,61
Becken	0,90	1,31
Humeri mit Scapulae	0,65	0,62
Femora	0,66	0,73

Tab. 15: Vergleich 75%-Perzentile  $qPET_{mean}$  physiologisches Skelett vs. 25%-Perzentile  $qPET_{mean}$  mit Skelettbefall, getrennt nach Skelettregionen

Der Median des Leber-SUV<sub>mean</sub> der 39 Datensätze lag bei 1,63 (Range 0,46 - 2,82). Betrachtet man den Median des Leber-SUV<sub>mean</sub> der 18 Patienten mit physiologischer Nuklidbelegung des Skeletts, so lag dieser ebenfalls bei 1,63 (Range 0,46 - 2,46). Von besonderer Bedeutung erscheint der Nachweis, dass auch bei sehr hoher oder niedriger Nuklidbelegung der Leber eine Normalisierung durch die Bildung von  $qPET$ -Werten gelingt. Hierzu wurden die Mediane mit Range der SUV<sub>mean</sub> der Skelettregionen bei 14 Patienten mit mittlerer Aktivitätsbelegung innerhalb des Leberquaders (SUV<sub>mean</sub> 1,31–2,20) den beiden Patienten mit erhöhter (2,46; 2,24) und verminderter (0,46; 0,87) Nuklidbelegung innerhalb des Leberquaders gegenübergestellt (Tab. 16). Gleichläufig zur Aktivitätsbelegung des Leberquaders erkennt man in den Stichproben eine deutliche Erhöhung bzw. Verminderung der Aktivitätsbelegung des Skeletts.

Region	Median und Range des SUV <sub>mean</sub> der Skelettregion bei		
	mittlerem Leber-SUV	vermindertem Leber-SUV	erhöhtem Leber-SUV
Wirbelsäule	1,42 (0,69 - 2,41)	0,82 (0,45 - 1,21)	3,53 (2,48 - 4,58)
Rippenthorax	0,90 (0,60 - 1,22)	0,46 (0,30 - 0,62)	2,06 (1,20 - 2,93)
Becken	1,23 (0,87 - 2,35)	0,77 (0,41 - 1,13)	2,79 (1,53 - 4,06)
Humeri mit Scapulae	0,82 (0,51 - 1,23)	0,54 (0,31 - 0,77)	1,70 (0,88 - 2,51)
Femora	0,77 (0,46 - 1,61)	0,42 (0,26 - 0,58)	1,74 (0,79 - 2,70)

Tab. 16: Median und Range der Aktivitätsbelegung der Skelettregionen, getrennt nach Nuklidbelegung des Leberquaders

Nach Normierung auf den Uptake der Leber zeigt sich eine Annäherung der  $qPET_{mean}$ -Werte zwischen diesen drei Gruppen (Tab. 17). Zudem sind alle Werte der Skelettabschnitte bei verminderter und erhöhter Aktivitätsbelegung des Leberquaders im Range der Werte bei mittlerer Aktivitätsbelegung des Leberquaders enthalten. Zusammenfassend lässt sich formulieren, dass nach einer Normierung weiterhin

Region	Median und Range der $qPET$ -Werte der Skelettregion bei		
	mittlerem Leber-SUV	vermindertem Leber-SUV	erhöhtem Leber-SUV
Wirbelsäule	0,83 (0,38 - 1,51)	1,18 (0,97 - 1,39)	0,86 (0,76 - 0,97)
Rippenthorax	0,54 (0,42 - 0,90)	0,68 (0,65 - 0,72)	0,48 (0,47 - 0,49)
Becken	0,79 (0,48 - 1,48)	1,10 (0,89 - 1,30)	0,64 (0,60 - 0,67)
Humeri mit Skapulae	0,50 (0,30 - 0,94)	0,78 (0,67 - 0,88)	0,38 (0,34 - 0,42)
Femora	0,48 (0,31 - 1,01)	0,62 (0,57 - 0,66)	0,38 (0,31 - 0,45)

Tab. 17: Median und Range der  $qPET$ -Werte der Skelettregionen, getrennt nach Nuklidbelegung des Leberquaders

Unterschiede in der Aktivitätsbelegung zwischen den einzelnen Skelettregionen bestehen, welche aber deutlich geringer ausfallen als ohne Normierung. In der untersuchten Kohorte von elf Patienten mit Skelettbefall (siehe Kapitel 6.2.3 Unterpunkt a)) befinden sich mehr als 80 Prozent der Läsionen in Wirbelsäule und Beckenskelett und diese beiden Regionen zeigen ähnliche physiologische Aktivitätsbelegungen. Ein einheitlicher Cut-Off für das gesamte Skelett sollte daher sicher Herde in diesen beiden Skelettregionen und damit mehr als 80 Prozent aller Skelettläsionen erfassen und dafür ggf. falsch-negative in anderen Regionen (Humeri, Femora, Rippenthorax) in Kauf nehmen.

#### **6.2.4 Ansatz IV: Vergleich der Aktivitätsbelegung von Wirbelkörpern und Wirbelkörperläsionen**

- c) Physiologischer Uptake eines Wirbelkörpers als Referenz für eine Unterscheidung zwischen physiologischem und pathologischem Skelettuptake

In Tabelle 18 wurden verschiedene Kennwerte zur näheren Charakterisierung der physiologischen Nuklidbelegung eines Wirbelkörpers erfasst. Der Median der nicht auf den Leberuptake normierten  $SUV_{mean}$  der 20 physiologischen Wirbelkörper betrug 2,19 mit einem Minimum von 1,35 und einem Maximum von 3,79. Hierbei konnte LWK 5 nur in fünf Fällen verwendet werden, da in den übrigen Fällen eine Hodgkinmanifestation vorlag. Eine Normierung auf den Leberuptake wurde hier nicht durchgeführt, da der physiologische Wirbelkörper selbst als Referenzregion zur Beurteilung von Wirbelkörperläsionen dienen soll. Die Standardabweichung des Durchschnittswertes des  $SUV_{mean}$  (in der Tabelle als  $SD_{WK}$  bezeichnet) betrug im Median 0,25 und zeigt, dass ein als physiologisch eingestufte Wirbelkörper einen relativ homogenen Nuklidbesatz aufweist.

Der nicht auf die Leber normierte Median der  $SUV_{peak}$ -Werte betrug 2,91 mit einem Minimum von 1,77 und einem Maximum von 4,66. Das zeigt, dass einzelne Voxel die durchschnittliche Aktivitätsbelegung des physiologischen Wirbelkörpers deutlich überschreiten. Diese dürfen nicht als kleine Hodgkinmanifestationen fehlgedeutet werden.

Physiologischer Referenzwirbelkörper	$SUV_{peak}$	$SUV_{mean}$	$SD_{WK}$
Mittelwert	2,90	2,34	0,28
Median	2,91	2,19	0,25
Minimum	1,77	1,35	0,13
Maximum	4,66	3,79	0,63

Tab. 18: Kennwerte bei 20 physiologischen Wirbelkörpern

#### d) Befallsläsionen

Im Gegensatz dazu betrug der Mittelwert des  $SUV_{peak}$  der 157 befallenen Wirbelkörper 4,88 mit einer Standardabweichung von 2,44; einem Minimum von 1,17 und einem Maximum von 16,86.

Die große Spanne des  $SUV_{peak}$  gibt einen Hinweis darauf, dass Läsionen innerhalb des Wirbelkörpers Unterschiede aufweisen und keine homogene Gruppe darstellen. Daher wurde sich erneut des Instrumentes der Normierung bedient: Hierzu wurden die  $SUV_{peak}$ -Werte der Läsionen auf den  $SUV_{peak}$  des jeweiligen physiologischen Referenzwirbelkörpers bezogen. Es wurde ein Mittelwert von 1,64 mit einer



Standardabweichung von 0,65 und einem Minimum von 0,65 bzw. einem Maximum von 6,64 gemessen.

Das mittels des Werkzeuges „RegionGrow“ (siehe Methodikteil 5.2.4 a)) vermessene, visuell kontrollierte, metabolische Volumen der Wirbelkörperläsionen betrug im Mittel 5,77 ml bei einem Minimum von 0,20 ml und einem Maximum von 35,01 ml. Im letzten Fall wies der gesamte Wirbelkörper eine durch das Hodgkin-Lymphom bedingte Mehranreicherung des Radiotracers auf. Diese Spannweite zeigt das breite Spektrum der Ausprägung eines Wirbelkörperbefalls. Große Läsionen von mehreren Millimetern Volumen sind wahrscheinlich sicher als solche zu identifizieren. Schwierig zu detektieren sind aber möglicherweise kleine Läsionen mit einem Volumen von deutlich weniger als 1 ml, weil diese Läsionen nur aus einzelnen Voxeln bestehen. Da auch einzelne Voxel physiologischer Wirbelkörper eine deutlich erhöhte Aktivitätsbelegung zeigen (vgl. Ergebnisse Kapitel 6.2.4 a) ist die Zuordnung dieser kleinen Läsionen bedeutsam für Sensitivität und Spezifität des Algorithmus.

Abbildung 45 zeigt den auf den physiologischen Referenz-SUV<sub>peak</sub> normierten Läsions-SUV<sub>peak</sub> gegenüber dem metabolischen Tumervolumen der Wirbelkörperläsionen. Hierbei zeigt sich eine schwache Tendenz zwischen Läsionsvolumen und der normierten Aktivitätsbelegung.

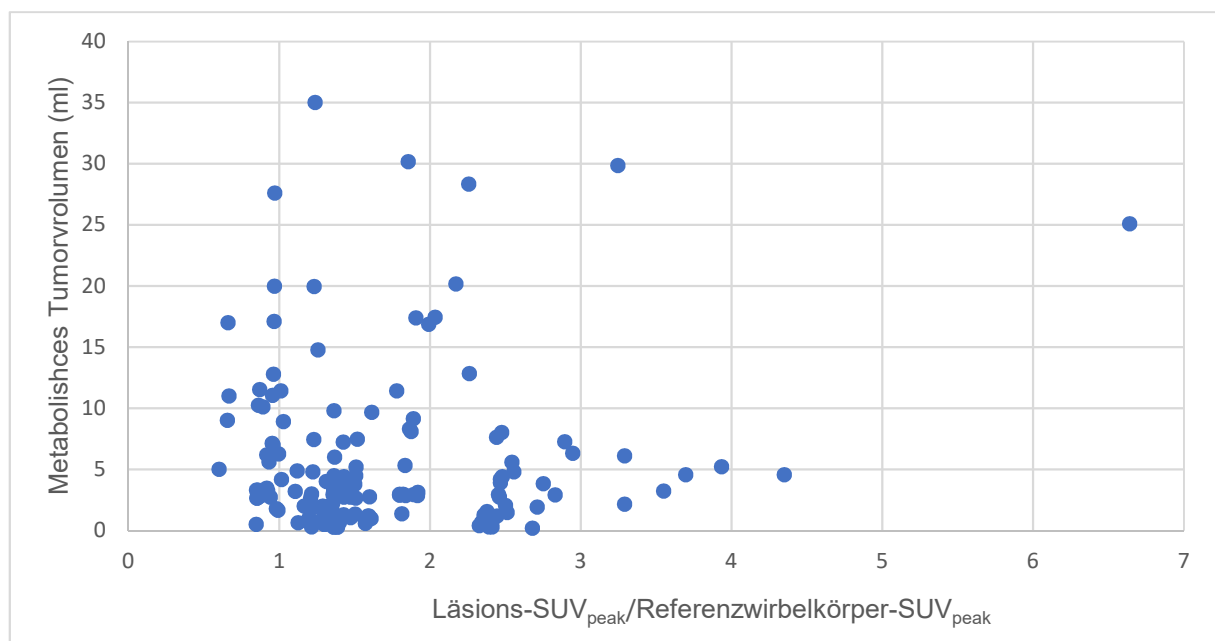


Abb. 45: Streudiagramm des normierten SUV<sub>peak</sub> gegenüber dem jeweiligen Läsionsvolumen

### **6.2.5 Ansatz V: Test verschiedener Standardabweichungen über dem Mittelwert der Aktivitätsbelegung des Gesamtskeletts bzw. des physiologischen Wirbelkörperuptakes als Grundlage für die programmgestützte Läsionendetektion**

Bei insgesamt elf Patienten mit visuell ausgeprägtem Skelettbefall fanden sich 152 Läsionen in den insgesamt 495 potenziellen Befallsregionen. 70 dieser 152 Läsionen imponierten bei der rein visuellen Bewertung flau-diffus, 82 Läsionen kräftig-fokal. Die übrigen 343 der 495 Regionen waren nicht befallen; anhand dieser wurde die Spezifität durch Anwendung unterschiedlicher Standardabweichungen in Bezug auf den Uptake des Referenzwirbels errechnet. Der Anteil der befallenen gegenüber den nicht befallenen Läsionen ist in Abbildung 46 dargestellt.

Hinsichtlich der Häufigkeit des Hodgkin-Befalls in den einzelnen Skelettabschnitten zeigten sich deutliche Unterschiede: Mit nur 3,9 Prozent Anteil befallener Halswirbelkörper ist die Halswirbelsäule (HWS) eher selten betroffen, die Brustwirbelsäule (BWS) schon häufiger mit 31,8 Prozent aller Brustwirbelkörper. Die Lendenwirbelsäule (LWS) und das Becken zeigen am häufigsten Läsionen mit einem Hodgkin-Befall in 50,9 Prozent bzw. 53,4 Prozent der Fälle. Sternum, Rippen und Femora sind in jeweils 27,3 Prozent, Humeri in 18,2 Prozent, die Scapulae in 13,6 Prozent aller Fälle betroffen.

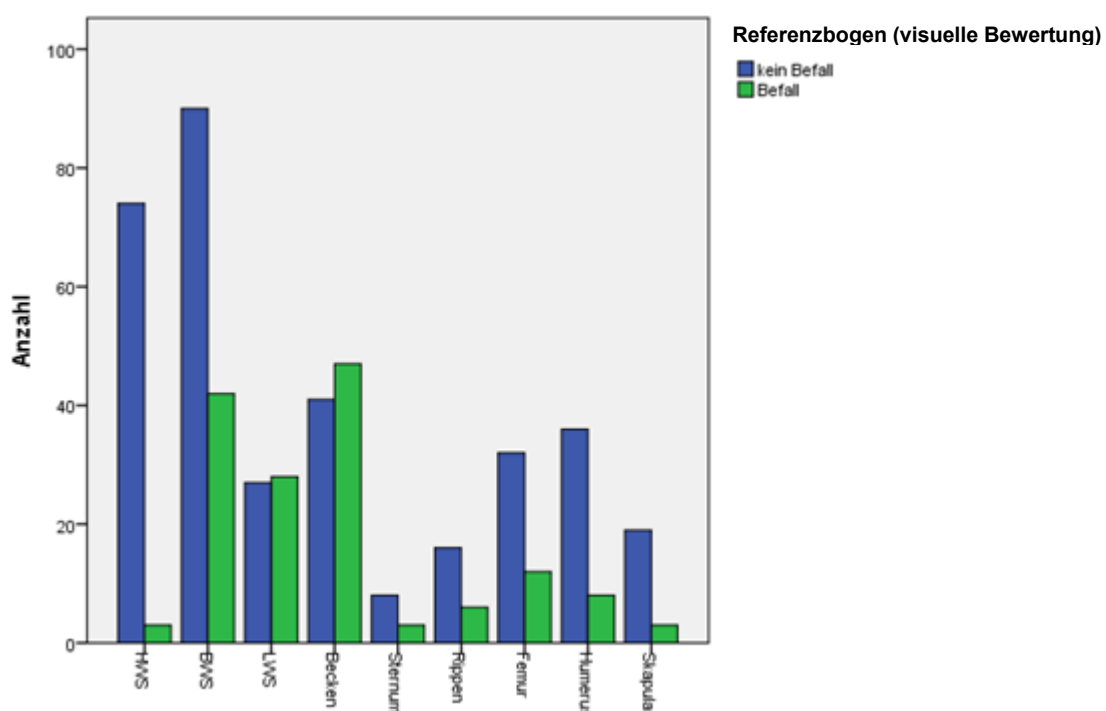


Abb. 46: Verteilung der Läsionen auf verschiedene Skelettregionen

Standardabweichung über Gesamtskelett	1,0	2,0	2,5	3,0	3,5
<b>Sensitivität (in %)</b>	98,6	86,8	59,2	44,1	38,8
- Für kräftig-fokale Läsionen (in %)	99,4	95,4	89,0	76,8	68,9
- Für flau-diffuse Läsionen (in %)	97,3	73,7	23,2	6,8	5,8
<b>Spezifität (in %)</b>	<b>28,2</b>	<b>69,2</b>	<b>89,8</b>	<b>97,7</b>	<b>100</b>

Tab. 19: Sensitivität und Spezifität bei Test verschiedener Standardabweichungen über der Aktivitätsbelegung der gesamten CT-Skelettmaske

Tabelle 19 zeigt, dass sich bei Verwendung eines Cut-Offs in Höhe einer Standardabweichung über dem  $SUV_{mean}$  innerhalb der CT-Skelettmaske mit 28,2 Prozent bzw. in Höhe von zwei Standardabweichungen mit 69,2 Prozent nur ungenügende Spezifitäten (rot) ergaben.

Im Gegensatz dazu wurde bei drei Standardabweichungen über dem  $SUV_{mean}$  mit 97,7 Prozent eine hohe Spezifität (grün) erzielt; 335 der 343 nicht befallenen Regionen wurden richtigerweise nicht angezeigt. Mit diesem Cut-Off wurden 76,8 Prozent der fokalen, aber nur 6,8 Prozent der flau-diffusen Läsionen korrekt (Goldstandard = visuelle Referenzbefundung) detektiert.

Abbildung 47 demonstriert (grüne Pfeile) am Übergang der Wirbelkörper BW 12 und LW 1 zum Pedikel zwei fokale Mehranreicherungen, welche typisch für Manifestationen des Hodgkin-Lymphoms sind. In LWK 4 ist hingegen eine flau-diffuse Läsion (roter Pfeil) markiert, welche nur fraglich einer Hodgkinmanifestation entspricht, möglicherweise visuell falsch-positiv und somit vernachlässigbar ist.

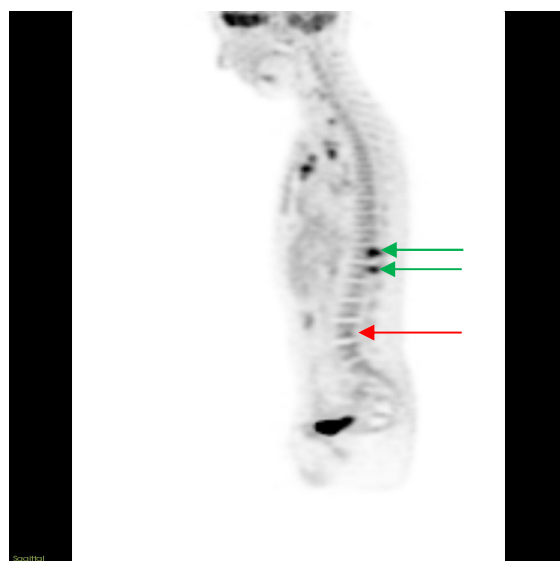


Abb. 47: Fokale Läsionen in LWK 1 und BWK 12, flau-diffuse Läsion in LWK 4

Zur genaueren Abgrenzung des Cut-Offs in Höhe von drei Standardabweichungen über dem mittleren Gesamtskelettuptake wurden zusätzlich Sensitivitäts- und Spezifitätsberechnungen in Höhe von zweieinhalb und dreieinhalb Standardabweichungen durchgeführt (Tab. 19): Bei dreieinhalb Standardabweichungen stieg die Spezifität auf 100 Prozent, wobei nur noch 5,8 Prozent der diffusen und 68,9 Prozent der fokalen Läsionen erkannt wurden. Bei zweieinhalb Standardabweichungen ergab sich eine immer noch sehr gute Spezifität von 89,8 Prozent, wobei 23,2 Prozent der diffusen Läsionen und 89,0 Prozent der fokalen Läsionen richtig erkannt wurden.

Tabelle 20 stellt die Ergebnisse des Algorithmus bei Anwendung verschiedener Standardabweichungen über der Aktivitätsbelegung des Referenzwirbelkörpers gegenüber. Bei zwei Standardabweichungen über dem Referenzwirbelkörper wurden 89,3 Prozent Spezifität erzielt, wobei 67,1 Prozent der diffusen Herde und 95,1 Prozent der fokalen Herde erkannt wurden. Zweieinhalb Standardabweichungen wiesen eine Spezifität von 95,4 Prozent auf, wobei 52,9 Prozent der diffusen Herde und 92,6 Prozent der fokalen Läsionen erkannt wurden. Drei Standardabweichungen zeigten die höchste Spezifität von 97,7 Prozent, wobei 40,0 Prozent der diffusen und 87,8 Prozent der fokalen Läsionen erkannt wurden.

<b>Standardabweichung über Wirbelkörper</b>	<b>2</b>	<b>2,5</b>	<b>3</b>
<b>Sensitivität (in %)</b>	82,2	74,3	65,8
- Für kräftig-fokale Läsionen	95,1	92,6	87,8
- Für flau-diffuse Läsionen	67,1	52,9	40
<b>Spezifität (in %)</b>	89,3	95,4	97,7

Tab. 20: Sensitivität und Spezifität bei Test verschiedener Standardabweichungen über der Aktivitätsbelegung eines physiologischen Wirbelkörpers

Zweieinhalb Standardabweichungen über der durchschnittlichen Aktivitätsbelegung des Wirbelkörpers wurden als optimaler Cut-Off des Algorithmus identifiziert. Dabei wurde eine sehr hohe Spezifität und in Bezug auf die kräftig-fokalen Herde eine sehr hohe Sensitivität von jeweils mehr als 92 Prozent erzielt.

### **6.2.6 Überprüfung des als optimal getesteten Algorithmus an Patienten mit physiologischem Skelettaufnahme**

Aus den Ergebnissen in Kapitel 6.2.5 ergab sich der Grenzwert der Aktivitätsbelegung von 2,5 Standardabweichungen über der mittleren Aktivitätsbelegung des Referenzwirbelkörpers, der für den Algorithmus im Weiteren verwendet wird.

Aus den Ergebnissen in Kapitel 6.2.4, wo sich zeigte, dass die vier zusammenhängenden Voxel mit der höchsten Aktivitätsbelegung in physiologischen Wirbelkörpern teils die Aktivitätsbelegung von Voxeln in Wirbelkörpermanifestationen überschritten, wurden vier Voxel als untere Grenzgröße für den Algorithmus abgeleitet.

Hieraus wurde folgender Algorithmus abgeleitet, welcher zur Überprüfung der Spezifität weiter getestet wurde:

1.)  $SUV_{Läsion} > SUV_{mean} \text{ des Wirbelkörpers} + 2,5 \text{ SD}$  und 2.)  $Größe_{Läsion} > 4 \text{ Voxel}$

Hierzu wurde der Algorithmus bei 50 Patienten ohne in der Referenzbegutachtung angegebenen Skelettbefall angewandt.

Bei 25 Patienten wurden hierbei keinerlei „falsch-positive“ Läsionen detektiert. Somit beträgt die Spezifität des Algorithmus auf Patientenebene 50 Prozent.

<b>Region</b>	<b>Verhältnis falsch-positiver Läsionen zur Regionenzahl</b>	<b>Median-Vol. (ml)</b>	<b>Minimum (ml)</b>	<b>Maximum (ml)</b>
LWS	14/250 (5,6 %)	0,38	0,26	1,34
BWS	21/600 (3,5 %)	0,35	0,25	0,77
Becken	8/400 (2,0 %)	0,43	0,25	0,7
Femur	2/200 (1,0 %)	1,08	0,99	1,16
Humerus	3/200 (1,5 %)	3,14	0,4	3,22
SC-Gelenk	1/50 (2,0 %)	0,3	0,3	0,3

Tab. 21: Verhältnis der falsch-positiven Läsionen gegenüber der geprüften Regionenzahl und Volumen falsch-positiver Läsionen, getrennt nach Skelettregionen

Bei den übrigen 25 Patienten zeigte der Algorithmus 49 „falsch-positive“ Läsionen auf, deren Lokalisation in Tabelle 21 jeweils mit Median und Range des Läsionsvolumens angegeben ist. 35 dieser Läsionen (71,4 Prozent) befanden sich in der Wirbelsäule. Auch bei Betrachtung des Verhältnisses der falsch-positiven Läsionen gegenüber der Anzahl der untersuchten Regionen treten in der Lendenwirbelsäule (5,6 Prozent) und der Brustwirbelsäule (3,5 Prozent) am häufigsten falsch-positive Läsionen auf. Alle fünf Läsionen in den proximalen Extremitäten (Femora und Humeri) betrafen die Wachstumsfugen. Vier dieser Läsionen wurden bei einem einzigen Patienten erfasst und zeigten das größte Volumen aller detektierten, falsch-positiven Läsionen. In den Humerus-Epiphysenfugen betrug das Volumen 3,14 ml und 3,22 ml und in den Femur-Epiphysenfugen 0,99 ml und 1,16 ml. In Abbildung 48 werden in einer Koronalrekonstruktion drei der vier Läsionen bei diesem Patienten blau farblich hervorgehoben.

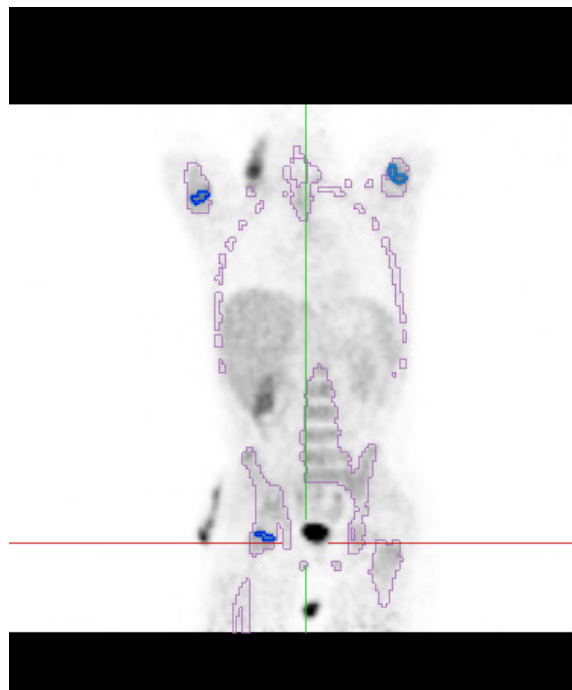


Abb. 48: Patient mit Mehranreicherungen in Epiphysenfugen von Humeri und Femora

Nur eine Läsion in einem Lendenwirbelkörper eines anderen Patienten war mit 1,34 ml vergleichbar groß, sonst wurden bei den übrigen falsch-positiven Läsionen Volumina  $\leq 0,77$  ml gemessen. Das Durchschnittsvolumen aller Läsionen betrug

0,56 ml +/- 0,59 ml. 34 von 49 falsch-positiven Läsionen (69,4 Prozent) wiesen ein Volumen von < 0,5 ml auf.

Bei insgesamt 14 der 25 Patienten mit falsch-positiven Befunden wurde nur eine falsch-positive Läsion gefunden, die übrigen elf Patienten wiesen bis zu fünf falsch-positive Läsionen auf.

	Durchschnitt	Standardabweichung	Minimum	Maximum
SUV <sub>mean</sub> des WK + 2,5 SD	3,21	1,16	1,09	7,65
SUV <sub>mean</sub> der falsch-positiven Läsionen	3,28	0,91	1,97	8,06
Überschreitung des Cut-Off durch die falsch-positiven Läsionen um (in %)	6,8	4,1	1,3	17,5

Tab. 22: Überschreitung der Aktivitätsbelegung der falsch-positiven Läsionen gegenüber dem Schwellenwert des Algorithmus

Tabelle 22 stellt den mittleren SUV-Grenzwert des Algorithmus (SUV<sub>mean</sub> des WK + 2,5 SD<sub>WK</sub>) dem mittleren SUV-Wert falsch-positiver Läsionen gegenüber. Vergleicht man auf Läsionsebene diese beiden Werte miteinander, so lag der mittlere SUV der Läsionen im Durchschnitt ca. 6,8 Prozent über dem mittleren SUV-Grenzwert, der sich nach Anwendung des Algorithmus ergab (Range: 1,3 bis 17,5 Prozent). Diese geringe Diskrepanz zwischen dem Grenzwert der Aktivitätsbelegung des Algorithmus und der Aktivitätsbelegung der falsch-positiven Läsionen ist ein weiteres Argument, dass dieser Grenzwert gut zur Diskrimination von physiologischen und pathologischen Skelettanteilen geeignet erscheint.

### **6.2.7 Überprüfung und Optimierung des Größenkriteriums zur Anwendung auf Datensätze mit unterschiedlicher Matrixgröße**

Aufgrund zahlreicher beteiligter Zentren in der EuroNet-PHL-C1-Studie mit unterschiedlichsten Geräten und Rekonstruktionsalgorithmen ergibt sich eine große

Bandbreite der Voxelgröße (vgl. Kapitel 5.2.2 und 5.2.7) von 5 bis 80 Voxeln pro 1 ml Volumen.

Bei 75 der 142 PET/CT ergaben 4 Voxel ein Volumen von 0,25 ml. Somit entsprachen die bisher vom Programm als Grenzgröße angewandten 4 Voxel bei den meisten Patienten etwa 0,25 ml. Dieses Volumen gleicht den kleinsten, visuell als Skelettbefall bewerteten Läsionen in der oben durchgeführten Größenanalyse von Wirbelsäulenläsionen (siehe Ansatz IV Ergebnisse Kapitel 6.2.4). Bei 43 der 142 Patienten setzen sich 0,25 ml aus weniger als 4 Voxeln und bei 24 der 142 Patienten aus mehr als 4 Voxeln zusammen.

Es wurde deshalb geprüft, ob sich die Verwendung eines standardisierten Größenkriteriums von 0,25 ml günstig auf die Treffsicherheit des Algorithmus im Vergleich zum zunächst gewählten Standard von 4 Voxeln auswirkt.

Es wurden hierzu 36 Patienten mit 256 erfassten Läsionen im Referenzbogen für Skelettbefall untersucht. Läsionen wurden vom Programm detektiert, wenn sie die unter Kapitel 6.2.5 und 6.2.6 als optimal erkannte Aktivitätsbelegung ( $SUV_{Läsion} > SUV_{mean} \text{ des Wirbelkörpers} + 2,5 \text{ SD}$ ) wie auch eine Größe von

**a) 4 Voxeln bzw.**

**b) 0,25 ml überschritten.**

Grenzgröße	Gesamtzahl	Als Läsion gewertet	Sensitivität (%)
4 Voxel	256	163	63,7
0,25 ml	256	183	71,5

Tab. 23: Vergleich der Sensitivität des Algorithmus bei Anwendung des Größenparameters 4 Voxel vs. 0,25 ml

Tabelle 23 zeigt, dass mit Überschreitung eines Größenkriteriums von 0,25 ml insgesamt 20 Läsionen mehr als bei einem Größenkriterium von 4 Voxeln detektiert worden sind. Damit ergab sich ein recht deutlicher Unterschied im Hinblick auf die Sensitivität (71,5 Prozent bei 0,25 ml versus 63,7 Prozent bei 4 Voxeln).

Somit konnte gezeigt werden, dass ein absolutes Volumen von 0,25 ml der absoluten Voxelzahl von 4 überlegen ist. Daher wird 0,25 ml als Grenzwert des Volumens für den Algorithmus übernommen.



### **6.3 Evaluation des experimentell ermittelten besten Programmalgorithmus zur Quantifizierung des Skelettbefalls in der FDG-PET an einer großen Patientenpopulation**

#### **1 Programmalgorithmus**

Als Ergebnis der Schlussfolgerungen aus Kapitel 6.2 wurde folgender Programmalgorithmus zur Anwendung auf eine große Patientenpopulation zugrunde gelegt:

1.  $SUV_{Läsion} > SUV_{mean}$  des Wirbelkörpers + 2,5 SD und
2.  $Größe_{Läsion} > 0,25$  ml

Wie im Abschnitt 5.3 beschrieben, wurden 142 pädiatrische Patienten aus der EuroNet-PHL-C1-Studie untersucht, die im **Initialstaging mindestens eine PET-positive Skelettläsion** aufwiesen. Insgesamt waren bei diesen 142 Patienten 1015 Läsionen im Referenzbogen für Skelettbefall verzeichnet worden. Somit hatte jeder Patient durchschnittlich 7,1 Skelettläsionen mit einer Range von 1 bis 30 Läsionen. Es wurden 131 Patienten retrospektiv und elf Patienten prospektiv untersucht.

#### **2 Art des Skelettbefalls**

Alle 142 Patienten wurden visuell nach der Art ihres Skelettbefalls in eine der beiden folgenden Gruppen eingeteilt:

##### **a) Tumorwachstum per continuitatem (s. g. E-Befall)**

Bei 6 der 142 Patienten (4,2 Prozent) lag ein E-Befall durch Tumorwachstum per continuitatem vor. Dieser betraf in allen sechs Fällen ausschließlich Rippen und Sternum. Diese Patienten wiesen keinen zusätzlichen hämatogenen Skelettbefall auf.

##### **b) Hämatogener Skelettbefall**

Bei den übrigen 136 der 142 Patienten (95,8 Prozent) lag ausschließlich ein hämatogener Skelettbefall vor. Diese Patienten wiesen keinen zusätzlichen E-Befall auf.

Ein gemischtes Befallsmuster aus hämatogenem Skelettbefall und Tumorwachstum per continuitatem in das Skelett lag somit bei den hier untersuchten Patienten nicht vor.

### **3 Verteilung und Aktivitätsbelegung der Referenzwirbelkörper mit Fallbeispiel**

Die verwendeten Referenzwirbelkörper wurden in Tabelle 24 zusammengestellt. Von 142 Patienten konnte in 96 Fällen (ca. 68 Prozent) der vierte Lendenwirbelkörper als Referenzregion verwendet werden. In 35 Fällen wurde ein anderer Lendenwirbelkörper als Referenz definiert (ca. 25 Prozent), da der vierte Lendenwirbelkörper befallen war. In elf Fällen (ca. 8 Prozent) musste ein Brustwirbelkörper als Referenz definiert werden, da die gesamte Lendenwirbelsäule befallen war.

Referenz-wirbelkörper	Anzahl	Anteil	Median SUV <sub>mean</sub>	Mittelwert SUV <sub>mean</sub>
<b>BWK 3</b>	1	0,7	2,04	2,04
<b>BWK 7</b>	1	0,7	1,90	1,90
<b>BWK 8</b>	1	0,7	1,30	1,30
<b>BWK 9</b>	1	0,7	2,98	2,98
<b>BWK 10</b>	3	2,1	2,47	2,59
<b>BWK 11</b>	3	2,1	1,81	1,83
<b>BWK 12</b>	1	0,7	2,89	2,89
<b>LWK 1</b>	4	2,8	1,81	1,81
<b>LWK 2</b>	3	2,1	1,96	2,03
<b>LWK 3</b>	23	16,2	1,93	1,93
<b>LWK 4</b>	96	67,6	1,95	1,84
<b>LWK 5</b>	5	3,5	1,79	1,80
<b>Gesamt</b>	<b>142</b>	<b>100</b>	<b>1,99</b>	<b>1,91</b>

Tab. 24: Referenzwirbelkörper (ohne Skelettbefall) für Suchalgorithmus

Der Mittelwert des  $SUV_{mean}$  des Referenzwirbelkörpers betrug 1,91 mit einer Standardabweichung von 0,66. Der Cut-Off-Wert, basierend auf der Formel  $SUV_{mean} + 2,5 SD$  lag im Mittel bei 3,36 ( $SD \pm 1,17$ ).

Die Abbildungen 49 und 50 stellen in Sagittal- und Koronalrekonstruktion das Ergebnis des angewandten Algorithmus exemplarisch dar. Grün markiert ist die CT-Skelettmaske. Der vierte Lendenwirbelkörper wies eine Hodgkinmanifestation auf, welche in der Koronalrekonstruktion durch einen roten Pfeil markiert wurde. Daher wurde als Referenzwirbelkörper der kranial davon liegende dritte Lendenwirbelkörper ohne Hodgkin-Befall als Referenzwirbel definiert, welcher türkis eingefärbt ist. Die vom Programm detektierten Läsionen sind blau umrandet.

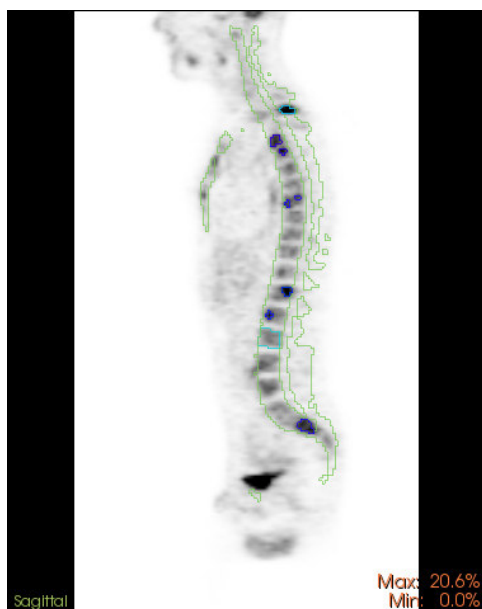


Abb. 49: Darstellung der Läsionen in der CT-Skelettmaske mit Referenzregion LWK 3 (Sagittalrekonstruktion)

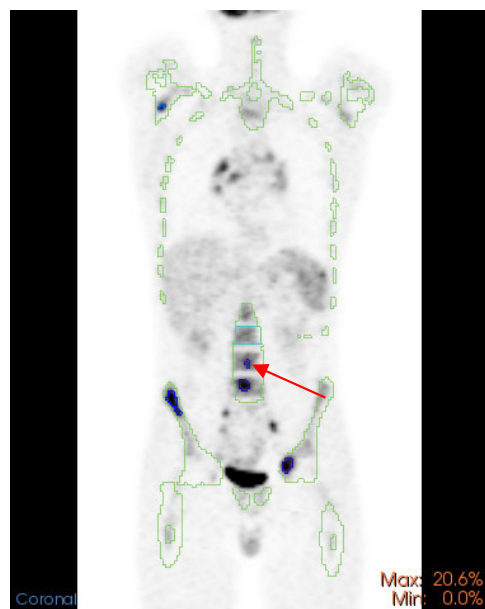


Abb. 50: Darstellung der Läsionen in der CT-Skelettmaske mit Referenzregion LWK 3 (Koronalrekonstruktion)

#### **4 Vom Algorithmus detektierte Aktivitätsmehrbelegungen**

Der Algorithmus detektierte bei den 142 Patienten insgesamt **1202** Aktivitätsmehrbelegungen im Skelett. Hiervon waren **964 richtig-positiv** und **238 falsch-positiv**.

### a) Richtig-positive Läsionen

964 vom Programm detektierten Läsionen konnten 774 Hodgkinmanifestationen aus der visuellen Referenzevaluation zugeordnet werden. Damit wurden durch das Programm in 190 Lokalisationen voneinander getrennte Läsionen detektiert, die in der Referenzbegutachtung als eine befallene Skelettregion dokumentiert wurden. Das betraf nahezu ausschließlich Hodgkinmanifestationen in Becken und Wirbelsäule. Das ist dadurch zu erklären, dass größere, inhomogene Läsionen aus Gruppen von Voxeln mit deutlicher pathologischer Aktivitätsmehrbelegung bestehen, zwischen denen Voxel liegen, welche eine geringere Aktivitätsbelegung zeigen und die Algorithmusparameter (siehe Kapitel 6.3, Unterpunkt 1) nicht überschreiten. Daher werden in einer inhomogenen Region mehrere Läsionen angezeigt.

Untenstehende Tabelle 25 listet die absolute Anzahl der Läsionen je Skelettabschnitt und den jeweiligen prozentualen Anteil an der gesamten Läsionszahl auf. Der überwiegende Anteil der Skelettläsionen befindet sich mit 83,6 Prozent in Wirbelsäule und Becken. Die verbliebenen 16,4 Prozent verteilen sich auf Rippen, Humerus, Femur, Scapula und Sternum.

Region	Anzahl Läsionen	Anteil an der Gesamtzahl der Läsionen (%)
Wirbelsäule	429	44,5
Becken	377	39,1
Rippen	23	2,4
Humerus	24	2,5
Femur	67	7,0
Scapula	10	1,0
Sternum	34	3,5
<b>Gesamt</b>	<b>964</b>	<b>100</b>

Tab. 25: Verteilung der richtig-positiven Läsionen auf verschiedene Skelettregionen

## b) Richtig-negative Regionen

Das Skelett jedes Patienten wurde in 45 Regionen unterteilt (siehe Methodik 5.2.5 a). Somit ergibt sich bei 142 Patienten eine Gesamtzahl von 6390 Regionen. In 1015 Lokalisationen lagen nach dem Referenzbogen für Skelettbefall Hodgkinmanifestationen vor. Somit waren bei den untersuchten 142 Patienten 5375 Regionen tumorfrei. In 5137 Regionen wurde vom Algorithmus auch keine Läsion angezeigt, somit waren diese Regionen als richtig-negativ erkannt worden.

## c) Falsch-positive Läsionen und Gründe für deren Detektion

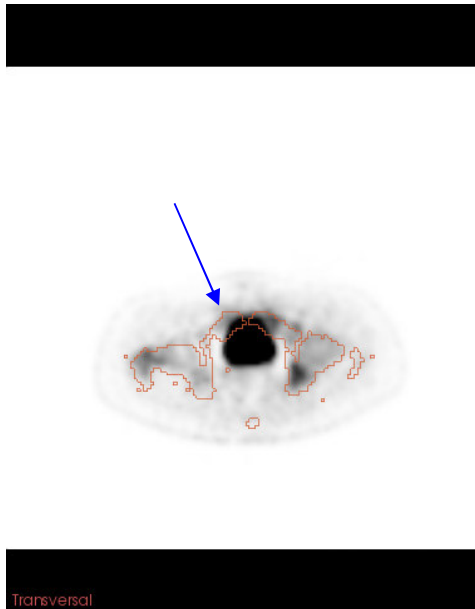
Insgesamt **238** nicht durch Skelettbefall des Hodgkin-Lymphoms verursachte Nuklidmehrbelegungen wurden durch das Programm in der CT-Skelettmaske erfasst (siehe Methodik Kapitel 5.1.2). Zum einen kam es in **134** Fällen durch **Spillover-Effekte** zur Projektion von **extraskelettaler Aktivitätsbelegung** auf das Skelett (siehe Methodik 5.3, Unterpunkt 6a). Zum anderen wurden **104** tatsächlich **im Skelett** gelegene Läsionen angezeigt, die nicht als Hodgkin-Läsionen in der visuellen Referenzbegutachtung bewertet wurden (siehe Methodik 5.3, Unterpunkt 6b).

### 1) Extraskelettale Aktivitätsbelegung mit Spillover-Effekt auf das Skelett

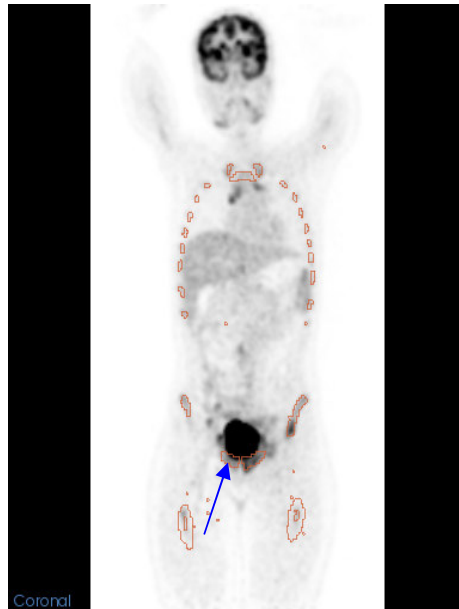
Die Ausscheidung des Radiopharmakons  $^{18}\text{F}$ -FDG erfolgt über die ableitenden Harnwege. Als Problemregion erwies sich hierbei das kleine Becken, wo sich Aktivität der den Schambeinen eng anliegenden Harnblase durch Spillover-Effekte auf das Skelett projiziert. Außerdem kann es durch unterschiedliche physiologische Füllungszustände der Blase zu einer Inkongruenz zwischen CT- und PET-Komponente kommen, da bei der nicht simultanen Aufnahme zwischen CT- und PET-Scan eine kurze Zeit verstreicht (Rodríguez-Vigil et al. 2006). Daraus resultiert bei Bildfusionierung die Projektion von Radiotracer aus der Blase auf den vorderen Beckenring. Daher detektiert der Algorithmus Läsionen in der CT-Skelettmaske, die in Wahrheit der mit Radiotracer gefüllten Harnblase entsprechen. Diese falsch-positiven Läsionen können leicht visuell erkannt und sofort verworfen werden, erfordern momentan aber zusätzliche manuelle Bearbeitungszeit.

Bei 62 der 142 Patienten bzw. 43,7 Prozent der Patienten wurde Blasenaktivität auf den vorderen Beckenring projiziert und als falsch-positive Läsionen vom Programm

detektiert (Abbildungen 51 und 52). Diese falsch-positiven Volumina betrugen bei den 62 Patienten im Mittel 3,98 ml mit einer Standardabweichung von 7,86 ml bei einer Range von 0,25 ml bis zu 57,48 ml.



**Abb. 51:** Blasenaktivität projiziert sich über die Schambeinäste (Transversalrekonstruktion)



**Abb. 52:** Blasenaktivität projiziert sich über die Schambeinäste (Koronalrekonstruktion)

Weiterhin traten Spillover-Effekte besonders an der Halswirbelsäule, den Rippen und dem Sternum durch die unmittelbare Nachbarschaft befallener Lymphknoten auf. Auch das Herz mit seinem teilweise ausgeprägten Glucosestoffwechsel konnte zu falsch-positiven Läsionen führen, sofern es Rippen oder Sternum direkt anlag. Da braunes Fettgewebe teilweise der Wirbelsäule direkt anliegt, projiziert es sich, wie in Abbildung 53 und 54 in Koronal- und Transversalrekonstruktion rot hinterlegt gezeigt, direkt auf die grün markierte CT-Skelettmaske.

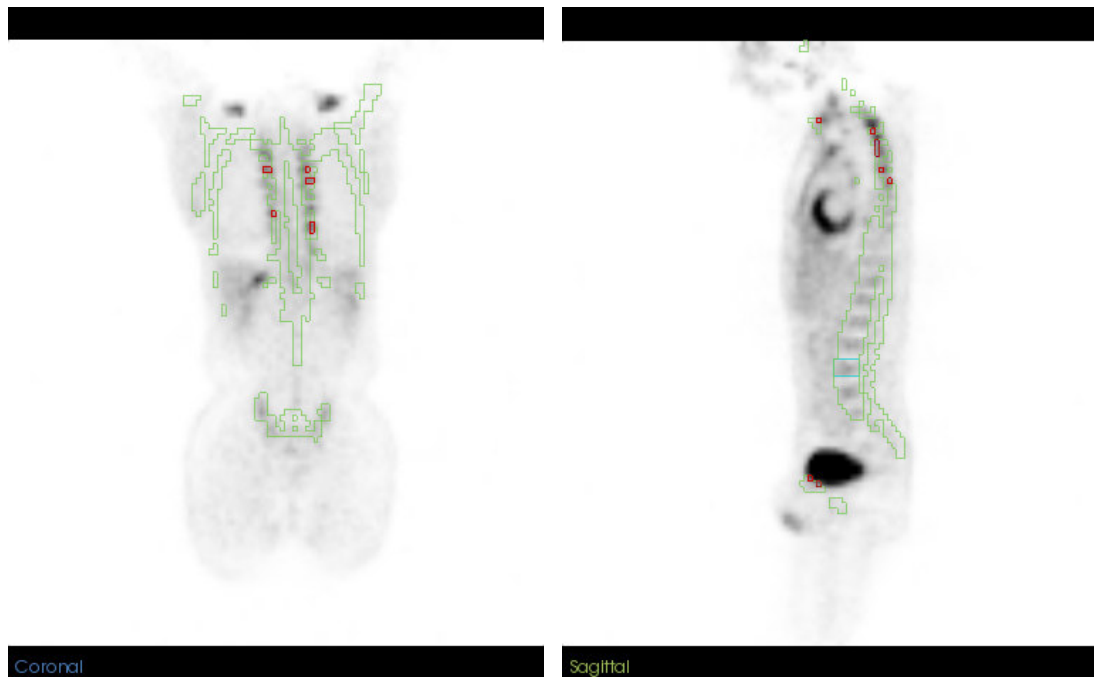


Abb. 53: Braunes Fettgewebe paravertebral in einer Koronalrekonstruktion dargestellt

Abb. 54: Braunes Fettgewebe paravertebral in einer Sagittalrekonstruktion dargestellt

Bei 72 von 142 Patienten bzw. einem Anteil von 50,7 Prozent der Patienten wurde ein durchschnittliches metabolisches Volumen von 4,1 ml bei einer Standardabweichung von 6,5 ml durch Spillover von Lymphknotenherden, Herz oder braunem Fettgewebe im Bereich des Thorax von der CT-Skelettmaske erfasst. Die Range reicht hierbei von 0,2 ml bis zu 42,6 ml.

## 2) Intraskelletale Aktivitätsmehrbelegungen, die im Referenzbogen für Skelettbefall nicht als Hodgkinmanifestationen gewertet wurden

Bei 51 der 142 Patienten (35,9 Prozent) wurden zusätzlich zu den im Rahmen der Referenzbegutachtung vermerkten Herdbefunden 104 intraskelletale Läsionen gefunden, welche nicht durch Spillover-Effekte umgebender Strukturen bedingt waren. Die Lokalisation dieser 104 zusätzlich erfassten Läsionen wurde in unten stehender Tabelle 26 geordnet nach Skelettregionen zusammengestellt. Hier fällt auf, dass der überwiegende Teil mit 72,1 Prozent in der Wirbelsäule liegt.

Region	Anzahl	Anteil an der Gesamtzahl der Läsionen (in %)
Wirbelsäule	75	72,1
Becken	13	12,5
Humerus	7	4,9
Femur	7	4,9
Klavikula	1	ca. 1
Sternum	1	ca. 1
<b>Gesamt</b>	<b>104</b>	<b>100</b>

Tab. 26: Lokalisation und Anteil der Skelettregionen an den zusätzlich detektierten Läsionen

Unten stehende Tabelle 27 stellt die Anzahl der zusätzlich detektierten Läsionen pro Patient dar. Bei 28 der 51 Patienten (54,9 Prozent) wurde nur eine Läsion zusätzlich zu den im Referenzbogen für Skelettbefall angegebenen Lokalisationen detektiert. Bei zehn Patienten wurden 2 Läsionen und bei 13 Patienten mehr als 3 Läsionen zusätzlich zum Ergebnis im Referenzevaluationsbogen detektiert.

Anzahl zusätzlicher intraskelettaler Läsionen pro Patient	Anzahl der Patienten	Prozentualer Anteil der Patienten
1	28	54,9
2	10	19,6
3	2	3,9
4	7	13,7
5	3	5,9
7	1	1,9
<b>Gesamt</b>	<b>51</b>	<b>100</b>

Tab. 27: Lokalisation und Anteil der zusätzlich detektierten Läsionen in verschiedenen Skelettregionen

Eine Reevaluation durch einen Facharzt für Nuklearmedizin steht noch aus. Sollten diese Läsionen schließlich nicht als durch das Hodgkin-Lymphom verursachte



Aktivitätsmehrbelegungen identifiziert werden, so müssen diese Läsionen tatsächlich als falsch-positiv eingestuft werden.

#### d) Falsch-negative Läsionen und Gründe für Nichtdetektion

Von den 1015 im Referenzbogen dokumentierten Läsionen wurden 241 durch den Algorithmus nicht detektiert.

Unten stehende Tabelle 28 schlüsselt die Gründe für die Nichtdetektion dieser 241 Läsionen nach den Parametern **zu geringe Größe** ( $< 0,25$  ml, siehe Methodik Kapitel 5.2.7), **zu geringe Aktivitätsbelegung** ( $< \text{SUV}_{\text{mean}} + 2,5 \text{ SD}$  des Referenzwirbelkörpers, siehe Methodik Kapitel 5.2.5) und **Nichterfassung** dieser skelettalen Läsionen in der **CT-Skelettmaske** (siehe Methodik Kapitel 5.1.2) auf.

Grund der fehlenden Detektion	Anzahl	Anteil (in %)
<b>Aktivitätsbelegung <math>&lt; \text{SUV}_{\text{mean}} + 2,5 \text{ SD}</math> des Referenz-WK</b>	129	53,5
<b>Volumen <math>&lt; 0,25</math> ml</b>	97	40,3
<b>Läsion liegt außerhalb der Skelettmaske</b>	15	6,2

Tab. 28: Häufigkeit der Gründe für die Nichtdetektion der falsch-negativen Läsionen

Aufgrund einer zu geringen Hounsfielddichte ( $< 160$  HU) der großen Knochenmarkhöhlen besonders von Femur und Humerus entstehen Fehlstellen in der CT-Skelettmaske, welche auch durch die Knochenfüllerfunktion nicht geschlossen werden können (siehe Ergebnisteil Kapitel 6.1.2). Ist eine Hodgkinmanifestation in solch einer Knochenmarkhöhle lokalisiert, so befindet sie sich somit außerhalb der CT-Skelettmaske. Um eine korrekte Quantifizierung des Ausmaßes des Knochenmarkbefalls zu gewährleisten, wurden alle 15 Läsionen mit Hilfe des „Region Grow“-Werkzeugs (siehe Methodikteil Kapitel 5.2.4) unter Anwendung des Algorithmusparameters ( $\text{SUV}_{\text{mean}} + 2,5 \text{ SD}$  des Referenz-WK) nachvermessen.

Unten stehende Tabelle 29 schlüsselt die Lokalisation der 241 nicht detektierten und somit falsch-negativen (FN) Läsionen auf. Bezogen auf die Gesamtzahl aller Läsionen wurden besonders viele Läsionen in der Wirbelsäule (46,1 Prozent) nicht detektiert, welche aber auch die häufigste befallene Skelettregion darstellt. Betrachtet man alle

im Erhebungsbogen für Skelettbefall angegebenen richtig-positiven und falsch-negativen Läsionen, ergeben sich insgesamt  $964+241=1205$  Läsionen. Vergleicht man nun die falsch-negativen mit den im Erhebungsbogen für Skelettbefall angegebenen Läsionen, so wird nur eine von fünf Läsionen (20,5 Prozent) in der am häufigsten befallenen Wirbelsäule durch den Algorithmus nicht detektiert.

Läsionen in Becken und Sternum werden nur in maximal 10 Prozent der Fälle nicht vom Algorithmus angezeigt.

Hingegen werden Läsionen im Schultergürtel und den Rippen besonders schlecht detektiert (in mehr als 60 Prozent der Fälle falsch-negativ).

Region	Anzahl FN	Anteil an allen FN (in %)	Gesamtzahl Positive	Verhältnis (in %) FN/Gesamtzahl
Wirbelsäule	111	46,1	540	20,5
Becken	32	13,2	409	7,8
Rippen	36	14,9	59	61,0
Humerus	16	6,6	40	40,0
Femur	25	10,3	92	27,2
Scapula	16	6,6	26	61,5
Sternum	2	0,8	36	5,6
<b>Gesamt</b>	<b>241</b>	<b>100</b>	<b>1205</b>	<b>20,0</b>

Tab. 29: Lokalisation und Anteil der falsch-negativen Läsionen in verschiedenen Skelettregionen

Bei zwölf Patienten, die alle zur retrospektiv getesteten größeren Kohorte gehörten, detektierte das Programm keinerlei Skelettherde. In der visuellen Referenzbegutachtung waren bis zu drei Läsionen aufgefallen (Tab. 30). Bei acht dieser zwölf Patienten bzw. 66,7 Prozent der Patienten lag aber nur eine einzige Läsion vor, bei jeweils zwei weiteren Patienten fanden sich zwei bzw. drei Läsionen, die nicht detektiert worden sind.

Anzahl nicht detektierter Skelettläsionen pro Patient	Anzahl der Patienten	Anteil dieser Patienten
1	8	66,7
2	2	16,7
3	2	16,7
<b>Gesamt</b>	<b>12</b>	<b>100</b>

Tab. 30: Anzahl der Läsionen bei Patienten mit Skelettbefall ohne durch den Algorithmus detektierte Herde

## 5 Kennwerte des Programmalgorithmus

### a) Sensitivität

Von den 1015 im Rahmen der visuellen Referenzbegutachtung ermittelten Läsionen wurden durch den Algorithmus 774 richtig erkannt, was einer Sensitivität von 76,3 Prozent auf Läsionsebene entspricht.

Hierbei wurden bei den 131 Patienten in der retrospektiven Testung 709 von 943 Läsionen richtig erkannt (Sensitivität von 75,2 Prozent) und bei den elf Patienten in der prospektiven Testung 65 von 72 Läsionen (Sensitivität von 90,2 Prozent).

Man kann den Algorithmus als Suchinstrument für Skelettbefall nutzen, damit zunächst eine Vorauswahl von Skelettmehrbelegungen angeboten wird. Diese können anschließend akzeptiert oder verworfen werden. Hierbei ist bedeutsam, dass mindestens eine Läsion bei jedem Patienten mit Skelettbefall detektiert wird.

Insgesamt wurde bei 130 von 142 Patienten mindestens eine Läsion detektiert, was einer Sensitivität von 91,5 Prozent auf Patientenebene entspricht.

### b) Spezifität

Der Referenzbogen für Skelettbefall stellt den Goldstandard zur Evaluation der Spezifität des vorgestellten Programmes dar (vgl. Methodik Kapitel 5.3, Unterpunkt 6). Es wird somit vorausgesetzt, dass in diesem Bogen alle tatsächlich vorhandenen Skelettläsionen auch erfasst wurden. Daraus folgt, dass alle zusätzlich zum Referenzbogen für Skelettbefall detektierten, intraskelettalen Läsionen als falsch-positiv bewertet werden.

Von den 5375 nicht befallenen Regionen wurde in 5137 Regionen auch keine Läsion vom Algorithmus angezeigt. Somit ergibt sich eine Spezifität auf Regionenebene von 95,6 Prozent.

Bei 91 der 142 Patienten wurde neben den im Referenzbogen für Skelettbefall vermerkten Hodgkinmanifestationen keine weitere, tatsächlich intraskelettale Läsion detektiert. Somit ergibt sich eine Spezifität auf Patientenebene von 64,1 Prozent.

## 6 Quantifizierung des Skelettbefalls

Das durchschnittliche Volumen der Summe aller richtig-positiven Läsionen eines Patienten betrug im Mittel 22,30 ml mit einer Standardabweichung von 35,56 ml und einem Maximum von 207,52 ml. Die Häufigkeitsverteilung ist in Form eines Histogramms in Abbildung 55 dargestellt. Bei 75 Patienten (52,8 Prozent) lag das Volumen bei weniger als 10 ml.

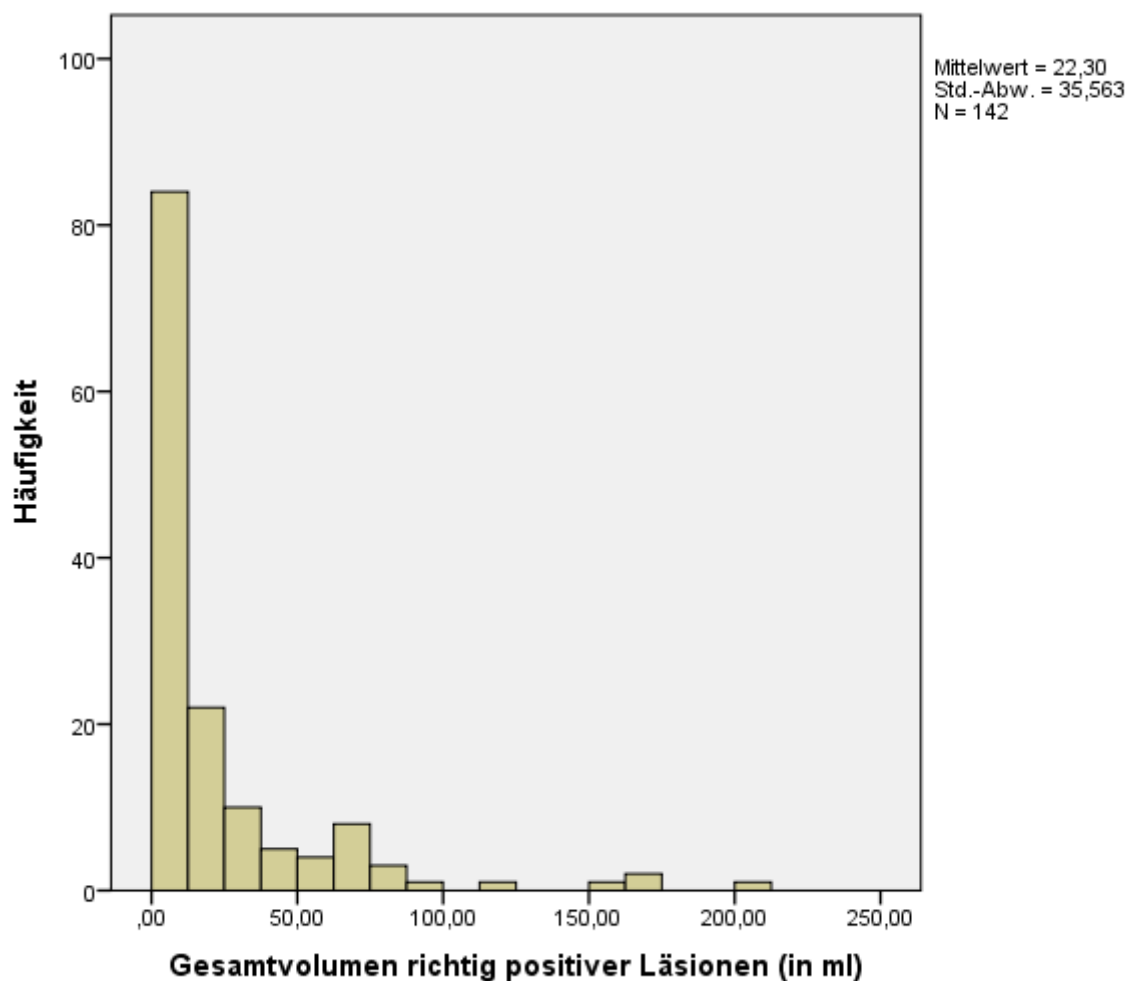


Abb. 55: Histogramm des Gesamtvolumens richtig-positiver Läsionen pro Patient

Der Anteil des Volumens der richtig-positiven Skelettläsionen am Gesamtskelettvolumen innerhalb der CT-Skelettmaske entsprach im Mittel 1,02 Prozent (Standardabweichung 1,90 Prozent, Maximum von 15,21 Prozent).

Abbildung 56 stellt ein Histogramm der Häufigkeitsverteilung dieses quantifizierten Skelettbefalls dar. Hierbei fällt auf, dass bei 91 Patienten (64,1 Prozent) und damit bei dem überwiegenden Anteil der Patienten weniger als ein Prozent des Skeletts befallen waren und nur bei vier Patienten (2,8 Prozent) mehr als fünf Prozent des Skeletts.

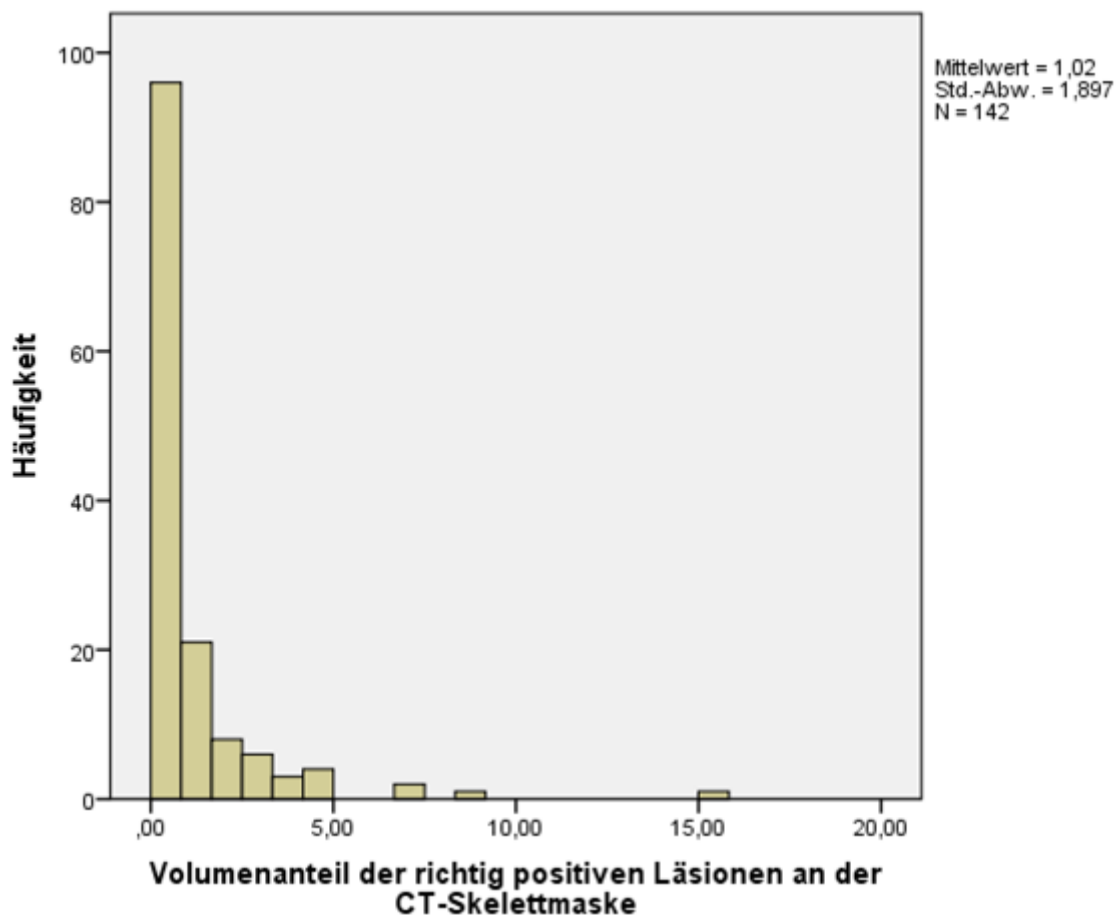
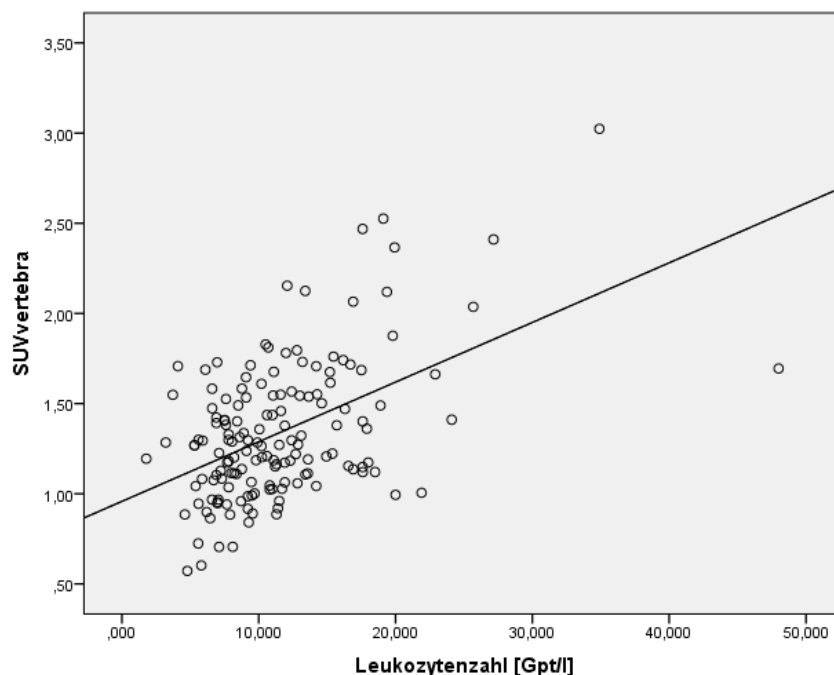


Abb. 56: Histogramm des Volumenanteils der richtig-positiven Läsionen an der CT-Skelettmaske

#### **6.4 Korrelation des durch den Algorithmus quantifizierten Skelettbefalls mit laborchemischen und klinischen Outcomedaten der Patienten**

Die Zahl der weißen Blutkörperchen wies gegenüber der Aktivitätsbelegung des Wirbelkörpers als potenziellen Surrogatmarker für den Grad der Knochenmarkaktivierung nur eine schwache Korrelation von 0,5 nach Pearson auf. Abbildung 57 stellt diesen Zusammenhang in einem Streudiagramm grafisch dar.



**Abb. 57: Streudiagramm zwischen Leukozytenzahl und SUV<sub>mean</sub> des Wirbelkörpers**

Für die übrigen laborchemischen bzw. klinischen Parameter aus der zentralen Datenbank (Hämoglobin, Gesamtprotein, Albumin, Leukozyten- und Lymphozytenzahl, Blutsenkungsgeschwindigkeit, Serumfibrinogen, Antikörperklassen IgG, IgA, IgM, Alpha-2-Globuline, Gammaglobuline, LDH, B-Symptome) wurden keine Korrelationen mit dem Ausmaß der pathologischen skelettalen Nuklidbelegung gefunden.

Es zeigte sich insbesondere keine Korrelation zwischen dem nodalen Tumervolumen und dem mittels Algorithmus ermittelten skelettalen Tumervolumen, wie das Streudiagramm der jeweils logarithmierten Volumina in unten stehender Abbildung 58 verdeutlicht.

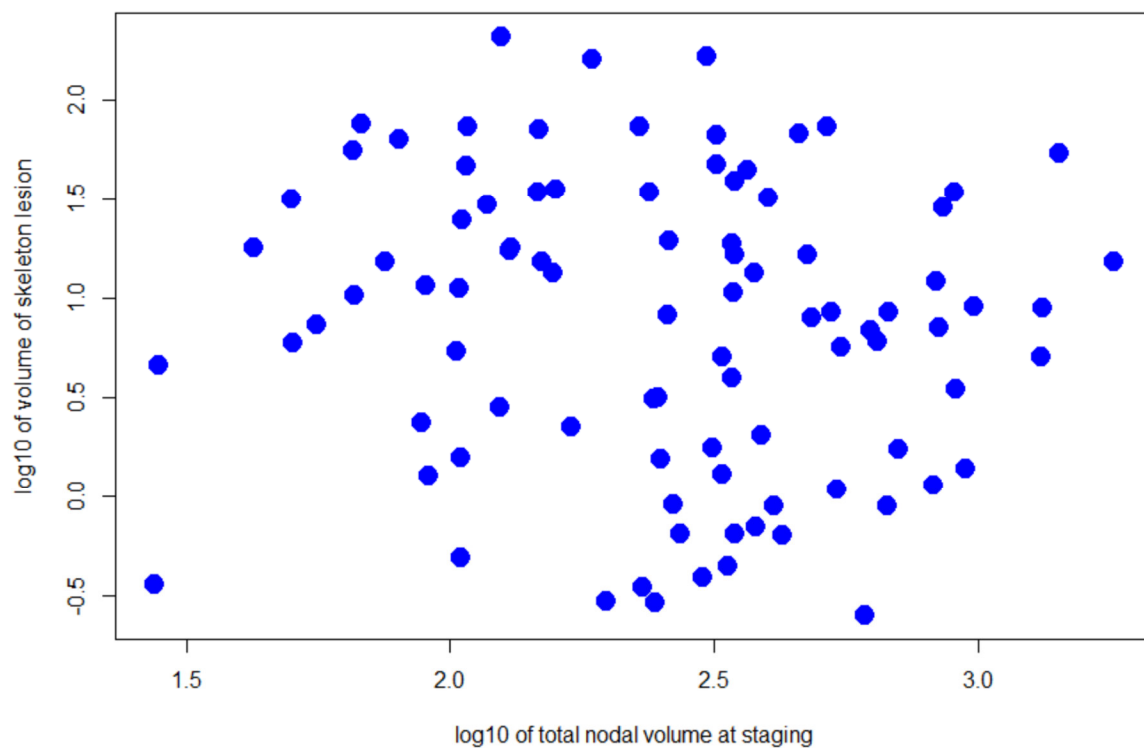
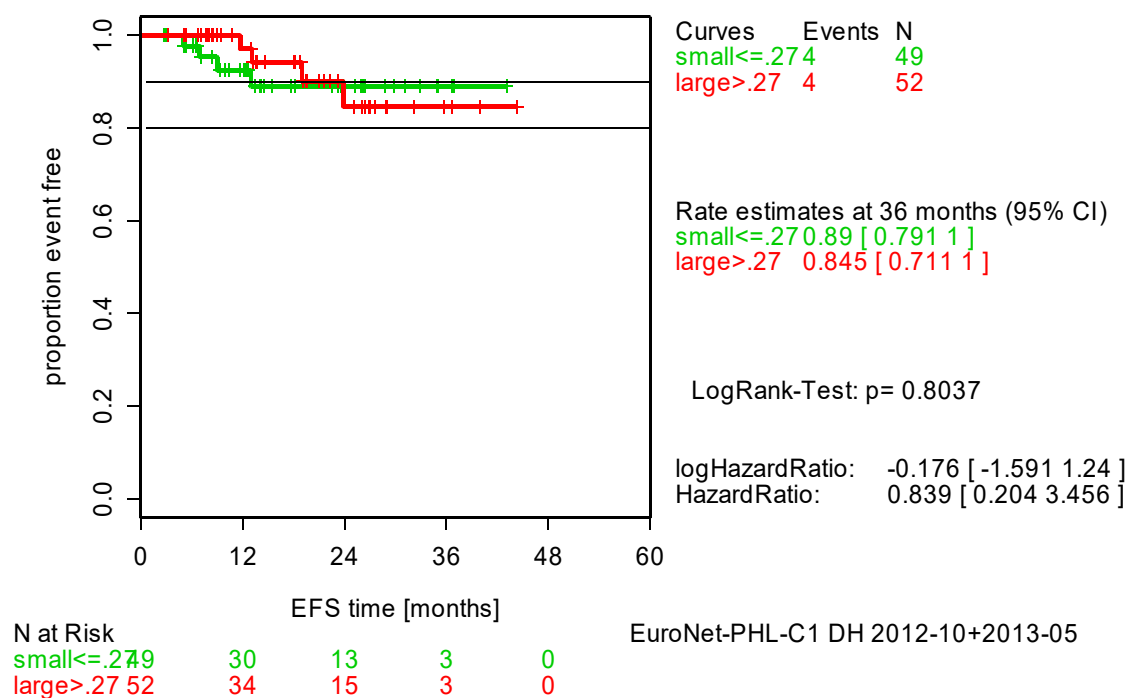


Abb. 58: Streudiagramm der logarithmierten Volumina der Lymphknotenherde und Skelettläsionen

Die Kaplan-Meier-Kurven (Abb. 59) demonstrieren, dass die Quantität des befallenen Skeletts keinen Einfluss auf das ereignisfreie Überleben hat. Sowohl in der Gruppe mit Skelettbefall von weniger als auch mehr als 0,27 Prozent der CT-Skelettmaske traten gleich viele Ereignisse (Rezidiv/Progress) im Beobachtungszeitraum von drei Jahren auf.

**EFS patients with bone/bone marrow involvement by lesion fraction**



**Abb. 59: Ereignisfreies Überleben der Patienten abhängig vom quantifizierten Skelettbefall**



## **7 Diskussion**

### **7.1 Extraktion des Skelettsystems aus dem PET und dabei auftretende methodenimmanente Probleme**

Bei der Extraktion des Skeletts aus dem PET-Datensatz mit Hilfe einer CT-Skelettmaske ließen sich verschiedene Probleme identifizieren, für die Lösungswege gefunden werden mussten.

Der im Programm vorimplementierte Schwellenwert von 160 HU zur Diskrimination zwischen Skelett und extraskelletalen Geweben wurde bei 115 und somit 81 Prozent der Patienten erfolgreich angewendet. Der in den Vorarbeiten von (Elsner et al. 2012) ermittelte, beste Schwellenwert von 160 HU bei der Erstellung einer CT-Skelettmaske konnte somit bestätigt werden.

Durch Modifikation des Hounsfielddichteschwellenwertes zwischen 130 bis 230 HU konnten auch bei den verbliebenen 27 Patienten verwendbare CT-Skelettmasken erstellt werden.

Noch vorhandene kleine Fehlstellen wurden durch die in Zusammenarbeit mit der Firma Hermes Medical Solutions, Stockholm (Schweden) implementierte Auffüllfunktion („Hole Filling Filter“) erfolgreich in die CT-Skelettmaske integriert. Nur in wenigen Fällen konnte verfettetes, gelbes Knochenmark durch seine niedrigere Dichte mit teils negativen HU-Werten dennoch nicht in diese Maske eingeschlossen werden. Das Hinzufügen zur CT-Skelettmaske erfolgte in diesen Fällen mittels einer Schneidefunktion („Crop VOI“-Funktion). Diese „Crop VOI“- Funktion wurde ebenso verwendet, um extraskelletale, sehr dichte Anteile, welche den Schwellenwert überschritten (z. B. intravasales oder intraenterales Kontrastmittel, Verkalkungen oder die metallische Patientenliege) wieder aus der CT-Skelettmaske zu entfernen.

Mithilfe der genannten Werkzeuge konnte die CT-Skelettmaske soweit verbessert werden, dass nach visueller Überprüfung eine zuverlässige Erfassung des Volumens aller im Datensatz abgebildeten Skelettabschnitte möglich war. Das wiederum ist für die korrekte Bewertung der späteren Läsionslast von Bedeutung gewesen.

(Sambuceti et al. 2012) postulieren, dass es nicht möglich ist, anhand eines festen Hounsfieldwertes knöcherne Strukturen im gesamten Körper zu erkennen, da verschiedene Knochensegmente unterschiedlicher Körperregionen voneinander differente Schwächungskoeffizienten zeigen. Anhand der großen Patientenzahl

konnte aber gerade dieser Ansatz bei einem Schwellenwert von 160 HU im Rahmen der vorliegenden Arbeit bestätigt werden.

Eine weitere Automatisierung oben genannter Schritte könnte in enger Zusammenarbeit mit der Firma Hermes Medical Solutions, Stockholm (Schweden) erfolgen, um die Vorarbeitszeit von etwa 15 bis 20 Minuten zur Extraktion des Skelettsystems aus dem Positronenemissionstomogramm weiter zu verringern.

Der in Kapitel 3.7 der Einleitung bereits vorgestellte Ansatz von (Sambuceti et al. 2012) bietet eine alternative Möglichkeit der Extraktion des Skeletts, genauer des intraossären Volumens aus einem FDG-PET/CT-Datensatz.

Folgende Punkte fallen beim Vergleich des in dieser Arbeit erprobten Programmes und dem von (Sambuceti et al. 2012) auf:

1. Die Sambuceti-Arbeitsgruppe extrahiert das Skelett durch eine schichtweise Analyse sehr präzise, sodass im Bereich der Wirbelsäule beispielsweise keine Bandscheiben integriert werden. Dieses Problem trat gelegentlich beim vorliegenden Programm auf (siehe Ergebnisse Kapitel 6.1.2). Bedeutsam ist die Integration extraskelettalen Gewebes in die CT-Skelettmaske, da hierdurch die Berechnung der Läsionslast in Bezug auf das Volumen der CT-Skelettmaske verfälscht wird. Allerdings erscheinen die geringen Volumina der Bandscheiben vernachlässigbar.
2. Für Lymphomstudien wird häufig intravenös und peroral verabreichtes, jodhaltiges Kontrastmittel appliziert, um Lymphknoten von Gefäßstrukturen oder Darmschlingen unterscheiden zu können (Rodríguez-Vigil et al. 2006). Aufgrund der Hounsfielddichte des Kontrastmittels bis zu 300 HU musste dieses in den für diese Arbeit vorliegenden Datensätzen mittels der „Crop VOI“-Funktion manuell aus der Maske entfernt werden.

Hingegen wurde in den PET/CT-Untersuchungen von Sambuceti kein jodhaltiges Kontrastmittel verabreicht, sodass offen bleibt, ob der Algorithmus der Sambuceti-Arbeitsgruppe Kontrastmittel von intraossärem Volumen auch dann sicher unterscheiden würde, wenn z. B. die der Brustwirbelsäule direkt anliegende Aorta in einer arteriellen Phase kontrastiert und damit deutlich

dichte angehoben wäre, wie in der vorliegenden Arbeit (siehe Methodik Kapitel 5.1.2 Unterpunkt 1).

3. Der Zeitaufwand des Benutzers zur Skelettextraktion mit dem in dieser Arbeit vorgestellten Programm betrug etwa 15 bis 20 Minuten, sodass für den Einsatz in der klinischen Routine in Zukunft weitere Optimierungen notwendig sind. Dem gegenüber müssen beim Programm der Sambuceti-Arbeitsgruppe nur wenige manuelle Anpassungen vorgenommen werden. Allerdings benötigt die anschließende automatische Extraktion des Skeletts etwa zwei Stunden. (Sambuceti et al. 2012) geben selbst an, dass das zwar für Forschungszwecke akzeptabel ist, für die klinische Routine aber wesentlich beschleunigt werden muss.
4. Osteolysen weisen geringere Hounsfielddichten als ein intaktes Skelett auf. In einigen Fällen wurden diese Läsionen nicht durch das hier getestete Programm in der CT-Skelettmaske erfasst. Da (Sambuceti et al. 2012) nur Datensätze mit physiologischen Skeletten ohne hämatologische Lymphommanifestationen bzw. Metastasen untersucht haben, wurden sie nicht mit solch einer Problematik konfrontiert. Da aber auch die Kortikalis bei Skelettläsionen mit betroffen sein kann, die (Sambuceti et al. 2012) als randbildende Kontur benötigen, ist zu erwarten, dass auch ihr Algorithmus Schwierigkeiten bei der Erfassung von Skelettanteilen mit großen Osteolysen zeigen würde.
5. Als wesentlicher Unterschied ist zu nennen, dass (Sambuceti et al. 2012) nur Datensätze eines einzigen Zentrums auswerteten.  
Im Gegensatz dazu erwies sich der im Rahmen der Multi-Center-Studie EuroNet-PHL-C1 in 18 europäischen Ländern mit über 130 Zentren in der vorliegenden Arbeit verwendete Algorithmus als robust und lieferte auch bei Datensätzen, die mit verschiedensten Scannersystemen und Bildakquisitionsprotokollen erstellt wurden, vergleichbare Ergebnisse. Eine erfolgreiche Anwendung des Sambuceti-Algorithmus auf Datensätze anderer Zentren muss hingegen erst erwiesen werden.

6. Der SUV des in dieser Arbeit verwandten Referenzwirbelkörpers in Lenden- oder Brustwirbelsäule betrug im Mittel 1,99 (+/- 0,66). (Sambuceti et al. 2012) ermittelten die Aktivitätsbelegung aus der gesamten Brust- und Lendenwirbelsäule ihrer Patienten mit einem mittleren SUV von 2,01. Auch wenn somit ein einzelner Wirbelkörper mit einem ganzen Abschnitt der Wirbelsäule verglichen wird, ergeben sich doch nahezu identische Mittelwerte der Aktivitätsbelegung. Das ist auch in Hinblick auf die Multizentrität der dieser Arbeit zugrundeliegenden Patientenkohorte bemerkenswert.

## **7.2 Experimentelle Ansätze zur Detektion von Skelettbefall in der PET nach erfolgter Programmoptimierung**

Nachdem die Optimierung der Extraktion des Skelettsystems aus dem PET erfolgt war, wurden verschiedene experimentelle Ansätze getestet, um eine Methode zu finden, die zuverlässig Skelettbefall detektieren kann. Hierbei zeigte sich, dass insbesondere die Unterscheidung eines ausgedehnten Knochenmarkbefalls gegenüber einer Knochenmarkaktivierung Probleme bereitete.

### **7.2.1 Ansatz I: Analyse des Nuklidspeicherverhaltens im extrahierten Skelett unter Normierung auf das Referenzorgan Leber**

1. Um die Vergleichbarkeit der Aktivitätsbelegung in der Multi-Center-Studie zu gewährleisten (verschiedene PET-Geräte, Rekonstruktionsalgorithmen, Menge applizierter Radiotracer, Zeitpunkt der Aufnahme u. a.) wurde eine Normierung gegenüber einem intrinsischen Hintergrund, der Leber, durchgeführt. Diese Vorgehensweise wurde von der Arbeitsgruppe von (Hasenclever et al. 2014) etabliert, die diese Normierung für vom Hodgkin-Lymphom befallene Lymphknoten in der Frühresponsebeurteilung verwendeten.

Bei Anwendung einer Normierung des Skeletts auf das Referenzorgan Leber konnte nachgewiesen werden, dass eine Homogenisierung der Aktivitätsbelegung der jeweiligen Gruppen mit unterschiedlicher skelettaler Lymphombeteiligung im Initialstaging gelingt.

2. Die Leber war bei allen in dieser Arbeit zu dieser Fragestellung untersuchten 61 Datensätzen nicht befallen, was die Seltenheit des Leberbefalls durch das Hodgkin-Lymphom von unter fünf Prozent aus den Ergebnissen der Arbeit von (Ilica et al. 2011) bestätigt. Durch die in dieser Arbeit erneut gezeigte homogene Aktivitätsbelegung der Leber mit einer Standardabweichung von im Mittel 0,17 eignet sich diese zur Verwendung als Referenzhintergrund.

3. Stichprobenartig konnte bei zehn Patienten in der Frühresponsebeurteilung ein Therapieansprechen im Skelett nach zwei Zyklen Chemotherapie durch eine deutliche Reduktion der Aktivitätsbelegung des gesamten Skeletts gezeigt werden. Allerdings konnte trotz Normierung auf den Leberhintergrund eine geringe residuelle Mehrbelegung nicht von einem vollständigen metabolischen Ansprechen ohne fokale Mehrbelegungen in der PET differenziert werden. Allerdings waren sieben der zehn Patienten lymphomfrei und nur drei der zehn Patienten wiesen noch bis zu drei Läsionen im ERA-PET auf, was bei den unterteilten 45 Skelettregionen weniger als sieben Prozent des Gesamtskeletts entspricht.

(Georgi et al. 2018) zeigten bei 426 skelettalen Läsionen im Initialstaging ein durchschnittliches Volumen von 3 ml und einen qPET-Wert von 2,87. Im ERA-PET waren noch 83 Läsionen mit einem qPET-Wert von 1,47 vorhanden. Die deutliche Reduktion der Aktivitätsbelegung und des Volumens erklärt, warum diese Läsionen in der Aktivitätsbelegung des gesamten Skeletts untergehen. Fraglich ist zudem, ob eine geringe Restaktivität des ursprünglichen Hodgkin-Befalls im Skelettsystem tatsächlich vitalem Tumorgewebe entspricht. Genauso gut könnte es sich bei der Restaktivität um Reparaturprozesse im initial befallenen Skelettabschnitt handeln, die keinesfalls eine Therapieintensivierung rechtfertigen würden, vgl. (Georgi et al. 2018). Um diese Fragestellung zu untersuchen, wäre eine Nachfolgestudie mit größeren Fallzahlen und Korrelationen zum rezidivfreien Überleben notwendig.

4. Keine sichere Abgrenzung gelang zwischen den Gruppen mit Knochenmarkbefall und Knochenmarkaktivierung im Initialstaging. Auf Probleme in der Differenzierbarkeit dieser beiden Gruppen hatten bereits (Elstrom et al. 2004) hingewiesen. In der vorliegenden Arbeit wies die Gruppe

mit diffuser Knochenmarkaktivierung eine so große Streubreite der Aktivitätsbelegung auf, dass sie sich mit den Werten aller weiteren Gruppen (physiologisches Skelett, Knochenmarkbefall, ERA) breit überlappte, siehe auch Ergebnisse Kapitel 6.2.1 und Abbildung 32. Das lässt sich am ehesten auf die unterschiedlich starken Ausprägungen der Knochenmarkaktivierung zurückführen.

### **7.2.2 Ansatz II: Histogrammanalyse**

Bei der Anfertigung und visuellen Analyse von Histogrammen konnte der Effekt der Normierung auf das Referenzorgan Leber analog zum vorherigen Ansatz erneut gezeigt werden. Visuell ließen sich die Kurven der Patienten mit physiologischem Skelett deutlich von denen mit ausgeprägtem Skelettbefall und Knochenmarkaktivierung unterscheiden. Es wäre somit möglich gewesen, über eine visuelle Analyse des Kurvenverlaufs diese Patientengruppen voneinander zu separieren. Allerdings ließen die Histogrammkurven keine visuelle Differenzierung zwischen ausgeprägtem Skelettbefall und diffuser Knochenmarkaktivierung zu. Die Hypothese, dass tumorbefallene Voxel eines Patienten eine relativ einheitliche Aktivitätsbelegung aufweisen, die sich dann im Histogramm als markanter Peak nachweisen lassen, bestätigte sich nicht.

Die von (Beare 2011) entwickelten histogrammbasierten Schwellenwerte (Analytics) erlauben vermutlich ebenfalls keine einfache Differenzierung von physiologischer gegenüber pathologischer Aktivitätsmehrbelegung. Es wurde stichprobenartig gezeigt, dass die Positionen der Analytics bereits in Histogrammen von physiologischen Skeletten extrem variable Werte annahmen. Aus diesem Grund wurde dieser Ansatz nicht weiterverfolgt.

Es wurde in den medizinischen Datenbanken keine Arbeit gefunden, welche über eine Histogrammanalyse des gesamten Skeletts Tumorerkrankungen detektieren kann, insbesondere nicht im PET/CT.

Es gibt Arbeiten, die die Histogrammanalyse in der Positronenemissionstomografie zur Charakterisierung der Heterogenität von malignen Tumoren eingesetzt haben. Heterogenität bei Tumoren ergibt sich durch das gleichzeitige Vorliegen von soliden Arealen mit verstärkter Angiogenese und Zellproliferationsrate gegenüber Arealen mit Hypoxie, Nekrose und zystischer Degeneration, Fibrose oder Einblutung.

Die Heterogenität der Tumoren kann man charakterisieren, indem man die sogenannte AUC-CSH-Methode (d. h. area under the curve - cumulative SUV histogram) anwendet. Hierzu wird die Tumoraläsion im axialen PET-Datensatz in der Schicht des größten Durchmessers charakterisiert. Die SUV-Werte werden in Abhängigkeit vom  $SUV_{max}$  in einem Histogramm aufgetragen. Es erfolgt eine Einteilung in 10 Prozent-Schritten, beginnend bei 0 Prozent bis 100 Prozent des  $SUV_{max}$ . Der daraus errechnete AUC-CSH-Wert ist umso niedriger, je heterogener die Läsion ist.

(Watabe et al. 2012) verglichen retrospektiv die AUC-CSH-Werte bei 21 großen intraabdominellen Tumorformationen im PET/CT, von denen sich histologisch neun als gastrointestinale Stromatumoren (GIST) und zwölf als maligne Lymphome herausgestellt hatten. Hierbei waren GIST (0,41) deutlich heterogener als Lymphome (0,64).

(Nakajo et al. 2015) wendeten die AUC-CSH-Methode bei 63 Patienten mit verschiedensten, PET-positiven, muskuloskelettalen Tumoren an, von denen 35 Läsionen benigne und 28 Läsionen maligne waren. Sie konnten zeigen, dass über die Heterogenität anhand eines AUC-CSH-Grenzwertes von 0,42 eine bessere Unterscheidung zwischen benignen und malignen Läsionen als bei Anwendung von  $SUV_{mean}$  und  $SUV_{max}$  gelingt, wobei maligne Tumoren deutlich heterogener als benigne Tumoren waren.

Diese Arbeiten verglichen die Heterogenität einzelner Läsionen miteinander, um einen Hinweis auf die Artdiagnose zu erhalten. Im Gegensatz dazu wurde in der vorliegenden Arbeit ein ganzes Organsystem (Skelett) bei pädiatrischen Hodgkin-Patienten im Initialstaging untersucht, welches auch ohne Tumorbefall eine inhomogene Untersuchungsregion darstellt, wie im Ergebnisteil unter 6.2.3 (Aktivitätsanalyse ausgewählter Skelettregionen mit und ohne Skelettbefall) herausgearbeitet wurde. In der vorliegenden Arbeit sollte über eine Histogrammanalyse in Bezug auf den Hintergrund der bereits heterogenen, physiologischen Aktivitätsbelegung eine zusätzliche pathologische

Aktivitätsmehrbelegung herausgefiltert werden, was nur bei ausgeprägtem Skelettbefall visuell gelang. Unmöglich wurde diese Diskrimination jedoch bei Patienten mit zusätzlicher Knochenmarkaktivierung, da diese sehr variable Ausprägungen und teils ähnliche Aktivitätsbelegungen wie ein Skelettbefall zeigen kann.

(Ganeshan et al. 2017) konnten erfolgsversprechende Ergebnisse bei Anwendung der Texturanalyse auf Lymphomläsionen im CT erzielen. Texturanalyse bedeutet hierbei die Heterogenität einer Untersuchungsregion zu charakterisieren. Hierzu wurden native Low-Dose-CT im Rahmen koregistrierter PET/CT-Datensätze im Initialstaging bei 45 Lymphompatienten verwendet. Es wurde ein gemischtes Kollektiv mit 18 Hodgkin-Patienten und 27 Patienten mit aggressiven Non-Hodgkin-Lymphomen untersucht, von den 30 Patienten nachfolgend ein negatives Interims-PET aufwiesen und somit eine bekanntermaßen bessere Prognose als solche mit positivem Interims-PET zeigen. Es wurde eine Filtrations-Histogramm-Methode angewendet, vgl. (Miles et al. 2013). Hierzu wird zunächst ein räumlicher Skalierungsfaktor festgelegt, der von 2 mm (fein) bis 6 mm (grob) reichen kann (s. g. Filtration). Anschließend wurde die Ausprägung des Gipfels des Histogramms (engl. Kurtosis) analysiert. Eine positive Wölbung des Gipfels geht hierbei definitionsgemäß über die Gauss-Normalverteilungskurve hinaus, eine negative verläuft darunter. (Ganeshan et al. 2017) konnten zeigen, dass bei Patienten mit negativem Interims-PET bei Anwendung eines räumlichen Skalierungsfaktors von 5 mm eine positive Wölbung des Histogramms im Initialstaging ein unabhängiger Prognosefaktor für das ereignisfreie Überleben ist. Die Gruppe der Patienten mit positivem Interims-PET (15 Patienten) war laut (Ganeshan et al. 2017) zu klein um die Auswirkung der Texturanalyse auf deren Prognose auszuwerten.

(Flechsigt et al. 2017) untersuchten die Hounsfieldddichte in der nativen Low-Dose-CT-Komponente im PET/CT bei 90 Lymphompatienten. Er untersuchte insgesamt 176 Lymphomläsionen, 76 Läsionen im Initialstaging und 100 Läsionen im Interimsstaging. 55 Patienten waren an einem Hodgkin-Lymphom und 35 Patienten an einem Non-Hodgkin-Lymphom erkrankt. Er nutzte einen semiautomatischen Algorithmus, um Größe und Dichte der Lymphknoten zu bestimmen. Beim Vergleich mit anderen Parametern wie  $SUV_{max}$  und Kurzachsendurchmesser der Lymphknoten zeigte sich in der Histogrammanalyse eine deutliche Korrelation mit höheren



Hounsfielddichtewerten und dem Vorliegen von Malignität. Er konnte 20 HU als Grenzwert für Malignität herausarbeiten, bei dessen Überschreiten 96,6 Prozent der Lymphknoten maligne waren. Einen ähnlich sensitiven Grenzwert für Benignität konnte er nicht ermitteln, bei einer Dichte von weniger als 5 HU erwiesen sich jedoch 80 Prozent der untersuchten Lymphknoten als benigne.

Bei den Arbeiten von (Ganeshan et al. 2017) und (Flehsig et al. 2017) wurden einzelne Läsionen im CT mittels Histogrammen analysiert. Histogramme eignen sich gut, um einzelne Läsionen und deren Textur zu analysieren und daraus Prognosefaktoren für ein Therapieansprechen abzuleiten.

Problematischer ist es aber mit Histogrammen ein größeres Areal wie ein ganzes Organsystem zu untersuchen, insbesondere wenn dieses eine so variable Aktivitätsbelegung wie das Skelett in der hier vorliegenden Arbeit zeigt. Wie bereits oben erörtert, zeigen schon physiologische Skelette unterschiedliche Aktivitätsbelegungen verschiedener Skelettregionen. Die Histogrammanalyse kann dann nicht das heterogene Aktivitätsbelegungsmuster einer disseminierten Knochenmarkaktivierung des Skeletts von einem multifokalen Skelettbefall unterscheiden.

### **7.2.3 Ansatz III: Aktivitätsanalyse ausgewählter Skelettregionen mit und ohne Lymphommanifestationen**

1. (Elstrom et al. 2004) erörtern bereits den höheren Anteil von rotem Knochenmark im Stammskelett. Passend dazu zeigt sich eine höhere Aktivitätsbelegung von Becken und Wirbelsäule gegenüber Thoraxskelett und Extremitäten bei der Analyse der sechs untersuchten Skelettregionen. Zusätzlich kann der Partialvolumeneffekt eine Rolle spielen, da dünne Knochen wie Rippen oder Scapula von diesem Effekt stärker betroffen sind.

Bedeutsam erscheint dieses Ergebnis für die Bestimmung eines einheitlichen Schwellenwertes zur Differenzierung zwischen pathologischer und physiologischer Aktivitätsbelegung, da physiologischerweise aktivitätsmehrbelegte Voxel in den entsprechenden knochenmarktragenden Regionen falsch-positiv durch einen Algorithmus angezeigt würden, der sich auf Regionen mit niedriger Aktivitätsbelegung bezieht.

Denkbar waren nach diesem Ergebnis die Verwendung:

- a) eines regionenspezifischen Schwellenwertes
- b) eines Mittelwertes aller Skelettregionen als Schwellenwert
- c) eines Schwellenwertes einer Region mit erhöhter physiologischer Aktivitätsbelegung, was wiederum eine Verringerung der Sensitivität in Regionen mit niedrigerer Aktivitätsbelegung bedeuten kann

2. Deutlich konnte in diesem Abschnitt herausgearbeitet werden, dass die Schädelkalotte aufgrund der Projektion von Hirngewebe mit seinem physiologisch hohen Glucosemetabolismus auf den Knochen (s. g. Spillovereffekt) nicht auswertbar ist.

Auch (Sambuceti et al. 2012) weisen darauf hin, dass der Spillover der Aktivitätsbelegung des Gehirns auf die Schädelkalotte keine zuverlässige Bestimmung des Traceruptakes in dieser Region erlaubt.

Aufgrund des sehr seltenen Befalls des Schädels beim Hodgkin-Lymphom erscheint dieser Umstand aber von zu vernachlässigender klinischer Relevanz. In einer von (Weiler-Sagie et al. 2014) untersuchten Patientenkohorte von 73 Patienten war die Schädelkalotte bei elf Prozent der Patienten befallen, wohingegen in der untersuchten Patientenkohorte der EuroNet-PHL-C1-Studie mit 142 Patienten die Schädelkalotte in keinem Fall eine Lymphombeteiligung aufwies.

Zudem war die Schädelkalotte in aller Regel nur unvollständig in den vorliegenden Datensätzen abgebildet. Daher wurde der Schädel durch Anwendung der „Crop VOI“-Funktion standardmäßig aus der CT-Skelettmaske entfernt. Zu gleichem Ergebnis wie die vorliegende Arbeit kommen auch Georgi et al. und Purz et al. (persönliche Kommunikation mit den Autoren).

#### **7.2.4 Ansatz IV: Charakterisierung und Vergleich der Aktivitätsbelegungen physiologischer Wirbelkörper und Wirbelkörperläsionen**

Im Rahmen dieser Arbeit wurden bei 20 pädiatrischen Patienten mit Hodgkin-Lymphom 157 Wirbelsäulenläsionen hinsichtlich Größe und Aktivität untersucht,

welche mittels PET/CT und konkordanten Ergebnissen in MRT oder CT visuell detektiert worden waren.

Eine isolierte Analyse von Wirbelkörperläsionen mittels PET/CT beim Hodgkin-Lymphom wurde bisher nicht durchgeführt. Allerdings gibt es Arbeiten, die sich mit mittels PET/CT erfassten spinalen Metastasen solider Tumoren beschäftigen:

(Torigian et al. 2011) analysierten halbautomatisch 31 spinale Skelettmetastasen im FDG-PET/CT bei 16 erwachsenen Patienten mit neun verschiedenen Primärtumorentitäten mittels einer ROVER (Region of interest (ROI) visualization, evaluation and image registration) genannten Software der Firma ABX aus Radeberg. Die größte Subgruppe in dieser Studie waren sieben Patienten mit Lungenkrebs. Es wurde kein Patient mit einer Lymphomerkrankung eingeschlossen.

(Torigian et al. 2011) untersuchten hierbei die Größe der spinalen Wirbelkörpermetastasen mittels MRT und PET.

Zur **morphologischen Größenbestimmung** verwendeten (Torigian et al. 2011) eine im maximalen Abstand von 45 Tagen zur PET/CT durchgeführte **MRT** ohne zwischenzeitliche Therapie. In einer sagittalen, fettgesättigten, T2-gewichteten Sequenz wurden die signalangehobenen Metastasen manuell mit der Computermouse umfahren.

In der **PET**-Komponente wurden die Metastasen zunächst manuell mit einem Kreis in einem Abstand von zwei bis vier Voxeln zur eigentlichen Läsion umrandet und der  $SUV_{max}$  erfasst. Anschließend wurde die Software ROVER mit einem Schwellenwert von 40 Prozent dieses  $SUV_{max}$  und einem Minimalvolumen von 1 ml angewandt, um das metabolisch aktive Volumen zu bestimmen. Neben dem  $SUV_{max}$  wurde auch der  $SUV_{mean}$  ermittelt.

(Torigian et al. 2011) ermittelten eine Durchschnittsgröße des metabolisch aktiven Volumens der Skelettläsionen von  $3,63 \pm 2,40$  ml (Range 0,96 ml - 10,69 ml) mittels PET-Komponente und von  $4,02 \pm 3,04$  ml (Range 0,82 ml - 13,69 ml) mittels manueller Abgrenzung in der MRT.

Die im Rahmen der vorliegenden Arbeit visuell kontrollierten, mittels „RegionGrow“-Funktion semiautomatisch umrandeten (siehe Methodik Kapitel 5.2.4), metabolisch aktiven Anteile von Wirbelsäulenläsionen zeigten demgegenüber einen Mittelwert der

Volumina von  $5,77 \pm 4,83$  ml (0,20 ml - 35,01 ml). Damit überschreitet die Läsionsgröße der metabolisch aktiven Wirbelsäulenherde beim pädiatrischen Hodgkin-Lymphom die Größe der Wirbelsäulenmetastasen bei (Torigian et al. 2011) um etwa 50 Prozent. Eine Erklärungsmöglichkeit wäre der Umstand, dass bei Lymphomen und Leukämien im Gegensatz zu Metastasen solider Tumoren knochenmarktragende Strukturen wesentlich häufiger einen Befall aufweisen und die Läsionen deshalb auch größer sind, was in der hier vorliegenden Arbeit auch erstmalig quantifiziert wurde.

(Torigian et al. 2011) erhoben einen Median des  $SUV_{max}$  der Skelettläsionen von 4,72 mit einem Range von 3,93 bis 8,03.

In der vorliegenden Arbeit wurde im Gegensatz dazu ein stabilerer und weniger vom Partialvolumeneffekt abhängiger  $SUV_{peak}$  (siehe (Wahl et al. 2009) und (Hasenclever et al. 2014)) als Mittelwert aus den vier Voxeln mit der höchsten Aktivitätsbelegung berechnet. Es wurde dabei ein Median des  $SUV_{peak}$  von 4,88 mit einer Range von 1,17 bis 16,86 ermittelt. Trotz dieser Unterschiede in der Methodik des Vergleichs von Skelettmetastasen mit skelettalen Lymphommanifestationen und einer Single-Center-Studie mit einer Multi-Center-Studie stimmen die Werte der höchsten Aktivitätsbelegung nahezu überein.

#### **7.2.5 Test verschiedener Standardabweichungen über dem Mittelwert der Aktivitätsbelegung des Gesamtskeletts bzw. eines physiologischen Wirbelkörpers als Grundlage für die Läsionendetektion**

Folgender Algorithmus hat sich nach Testung verschiedener Standardabweichungen über dem Mittelwert der Aktivitätsbelegung des Gesamtskeletts wie auch eines physiologischen Wirbelkörpers als überlegen herausgestellt:

3.  $SUV_{Läsion} > SUV_{mean}$  des Wirbelkörpers + 2,5 SD und
4.  $Größe_{Läsion} > 0,25$  ml

1. Sensitivität und Spezifität des Algorithmus auf Läsionsebene (siehe Kapitel 3.2.5) haben sich bei Anwendung der Referenzregion **physiologischer Wirbelkörper** gegenüber der mittleren Aktivitätsbelegung der **CT-**

**Skelettmaske** und Spezifität innerhalb des PET-Datensatzes als überlegen herausgestellt.

Die geringere **Sensitivität (59,2 Prozent** bei  $SUV_{Läsion} > SUV_{mean}$  der CT-Skelettmaske + 2,5 SD) bei Verwendung der CT-Skelettmaske gegenüber dem Wirbelkörper (**74,3 Prozent**) als Referenzregion ist vor allem durch eine niedrigere durchschnittliche Aktivitätsbelegung bedingt. Das kann durch zwei Umstände bedingt sein:

- a) In der CT-Skelettmaske sind Skelettabschnitte (z. B. Extremitätenknochen) mit einem deutlich geringeren Anteil roten und höheren Anteil gelben Knochenmarks (s. g. Fettmark) im Vergleich zu Wirbelkörpern enthalten, siehe (Blebea et al. 2007). Daraus resultiert eine geringere Aktivitätsbelegung der gesamten Skelettmaske gegenüber einem Lendenwirbelkörper, da gelbes Knochenmark metabolisch weniger aktiv als rotes ist (Shreve et al. 1999).
- b) Die durch Anwendung des „Hole Filling Filter“ in die CT-Skelettmaske integrierten Zwischenwirbelscheiben verändern durch fehlende Aktivitätsbelegung die Aktivitätsbelegung der CT-Skelettmaske hin zu niedrigeren Werten. Allerdings sollte dieser Effekt durch die geringen Volumina von zu vernachlässigender Bedeutung sein, insbesondere im Vergleich zum Effekt des eingeschlossenen gelben Fettmarks wie oben beschrieben.
- c) Da die Skelettmaske aus heterogenen Skelettregionen besteht (vgl. Kapitel 6.2.3), ist die Standardabweichung höher als bei Verwendung eines homogenen Referenzwirbelkörpers mit einer geringen Standardabweichung (vgl. Kapitel 6.2.4), wobei eine Referenzregion mit geringerer Standardabweichung einen präziseren Schwellenwert liefert.

Die schlechtere **Spezifität (89,6 Prozent** bei  $SUV_{Läsion} > SUV_{mean}$  der **CT-Skelettmaske** + 2,5 SD) gegenüber **95,4 Prozent** bei Anwendung des **Wirbelkörpers** ist zum Teil dadurch zu erklären, dass bei Anwendung der gesamten CT-Skelettmaske die Aktivitätsbelegung durch die Läsionslast

beeinflusst wird. Der Schwellenwert des Algorithmus wird bei ausgeprägtem Skelettbefall durch zahlreiche Skelettherde so erhöht, dass im befallenen Skelett ( $SUV_{avg} = 1,30$ ) eine deutlich höhere Aktivitätsbelegung als im physiologischem Skelett ( $SUV_{avg} = 1,13$ ) vorliegt (siehe Tab. 2 in Kapitel 6.2.1)

2. Im Gegensatz zu den übrigen bis dahin getesteten Methoden (Normalisierung des Skeletts durch Relation zu einem Leberquader, Histogrammanalyse des Skeletts, separate Analyse nach Skelettregionen) gelang erstmalig eine praktikable **Unterscheidung zwischen Knochenmarkaktivierung und disseminierten Skelettherden**.

Das lässt sich darauf zurückführen, dass Wirbelkörper bei pädiatrischen Patienten einen hohen Anteil roten Knochenmarks enthalten. Eine Knochenmarkaktivierung verändert die Aktivitätsbelegung von Skelettanteilen mit rotem Knochenmark hin zu höheren Werten, so auch in diesem Wirbelkörper. Die übrigen durch die Knochenmarkaktivierung aktivitätsmehrbelegten Skelettabschnitte überschreiten nicht den auf diesen Referenzwirbelkörper bezogenen Schwellenwert im Gegensatz zu skelettalen Lymphomherden mit darüber hinaus gesteigerter Aktivitätsbelegung. Probleme bei der Detektion können nur dann auftreten, wenn Lymphomläsionen einen Uptake ähnlich der Knochenmarkaktivierung aufweisen. Diese würden dann jedoch auch bei der visuellen Analyse dem Nachweis entgehen.

3. Da beim Hodgkin-Lymphom die Aufrechterhaltung des malignen Geschehens einen ständigen Nachschub an neuen Blutzellen benötigt (Kurch et al. 2013b), wird das rote Knochenmark durch endogene hämatopoetische Wachstumsfaktoren zur Produktion von Blutzellen angeregt. Daraus ergibt sich eine konsekutiv erhöhte Aktivitätsbelegung des Knochenmarks, die teils schwierig von Skelettbefall abzugrenzen ist (Elstrom et al. 2004).

Lenden- und Brustwirbelsäule bestehen bei pädiatrischen Patienten nahezu vollständig aus rotem Knochenmark (Jung et al. 2000). Da die verwendeten Referenzwirbelkörper im Rahmen dieser Arbeit stets in Brust- oder Lendenwirbelsäule lagen und die Aktivitätsbelegung derselben ermittelt wurde, lag der Schluss nahe, dass sich das Ausmaß der Knochenmarkaktivierung

anhand der Aktivitätsbelegung dieser Wirbelkörper abschätzen lassen müsste, da Knochenmarkaktivierung die Aktivitätsbelegung des roten Knochenmarks steigert.

(Murata et al. 2006) haben die Eignung der Aktivitätsbelegung einzelner Wirbelkörper in der PET/CT-Untersuchung als Surrogatmarker für die Knochenmarkaktivierung bei 48 Patienten mit einem mittleren Alter von 64 Jahren untersucht. Sechs Patienten erhielten die PET/CT-Untersuchung aufgrund benigner Erkrankungen und 42 Patienten aufgrund maligner Erkrankungen. Sie werteten die Aktivitätsbelegung einer ca. 30 cm großen ROI aus, welche sie zum einen in den elften und zwölften Brustwirbelkörper und zum anderen in den ersten und zweiten Lendenwirbelkörper positionierten. Sie korrelierten diese Aktivitätsbelegung mit den verschiedenen Blutzellreihen wie auch mit der Hämoglobinkonzentration und dem Hämatokrit. Die Ergebnisse von (Murata et al. 2006) zeigten eine Korrelation der FDG-Anreicherung im Knochenmark mit der Leukozytenzahl im Blut, welche auch in unserer Arbeit beobachtet wurde.

Die von ihnen gefundene Korrelation von 0,59 bei einem Signifikanzniveau von  $p < 0,0001$  ist etwas stärker ausgeprägt als die in dieser Arbeit ermittelte von 0,5 bei einem Signifikanzniveau von  $p < 0,0001$ .

4. Wirbelkörper weisen gegenüber anderen Referenzregionen verschiedene Vorteile auf:
  - a) Ein einzelner Wirbelkörper kann sehr rasch visuell auf eine fokale Mehrspeicherung des Tracers durch Skelettbefall bzw. eine Minderspeicherung z. B. aufgrund eines Hämangioms geprüft werden, da nur wenige Schichten der PET/CT-Untersuchung bzw. ein Fokus einer Maximum-Intensitäts-Projektion (MIP) beurteilt werden müssen. Ein so in seiner Aktivitätsbelegung veränderter Wirbelkörper wäre als Referenzregion für den Algorithmus nicht geeignet, da der Schwellenwert verfälscht würde. In einem solchen Fall stünden jedoch angrenzende Wirbelkörper als geeignete alternative Referenzregionen zur Verfügung.

- b) Wirbelkörper weisen annähernd eine Zylinderform auf und lassen sich dadurch schnell und einfach manuell abgrenzen. Insbesondere die in den meisten Fällen (92,3 Prozent) verwendeten Lendenwirbelkörper sind zudem aufgrund ihres relativ großen Volumens (je nach Alter etwa 15 - 40 ml, ca. 1 Prozent der Skelettmasse) gut eingrenzbar (VOI-Technik). In 7,7 Prozent der Fälle musste bei visuell detektiertem Befall der gesamten Lendenwirbelsäule auf einen Brustwirbelkörper ausgewichen werden.
- c) Es konnte gezeigt werden, dass nicht erkrankte Wirbelkörper eine relativ homogene Aktivitätsverteilung aufweisen, da die Standardabweichung des SUV bei 192 analysierten Patienten (142 Patienten mit Skelettbefall und 50 Patienten ohne Skelettbefall) bei einem mittleren SUV von  $0,55 \pm 0,22$  lag.

### **7.3 Evaluation des experimentell gefundenen besten Programmalgorithmus zur Quantifizierung des Knochenmarkbefalls an einer großen Patientenpopulation**

#### **7.3.1 Sensitivität des Algorithmus**

Wenn das gesamte Skelett als Bewertungsgrundlage des Algorithmus genommen wird, so werden regionenspezifische Unterschiede verschiedener Skelettanteile gemittelt. Die Verwendung einer Region mit einer höheren physiologischen Aktivitätsbelegung hat den Vorteil, keine falsch-positiven Läsionen in Regionen mit geringerer Aktivitätsbelegung zu zeigen. Dahingegen werden kleine Läsionen oder Läsionen, die nur einen gering über dem physiologischen Skelett liegenden Uptake zeigen, womöglich nicht detektiert.

1. Bei 130 von 142 Patienten wurde mindestens eine visuell als Befallsläsion bestätigte Mehranreicherung im Skelett detektiert. Wertet man dies als Skelettbefall im Sinne eines Ann Arbor-Stadiums IV, so beträgt die Sensitivität des Programmes 91,5 Prozent auf Patientenebene.

Bei den verbliebenen zwölf Patienten, die alle zur retrospektiv getesteten größeren Kohorte gehörten, detektierte das Programm keinerlei Skelettherde. In der visuellen Referenzbegutachtung waren bei diesen Patienten bis zu drei



Läsionen markiert worden, wovon bei zehn dieser zwölf Patienten nur eine oder zwei Läsionen im Rahmen der Referenzbegutachtung als Skelettbefall dokumentiert worden waren.

Beachtet man, dass im neuen, seit September 2015 gültigen Studienprotokoll (EuroNet-PHL-C2-Studie; EudraCT-Nr. **2012-004053-88**) mindestens 3 Skelettläsionen in der FDG-PET zur Diagnosestellung des Knochenmarkbefalls und damit einem Upstaging ins Stadium IV erforderlich sind (Purz et al. 2011), so läge bei zehn der zwölf Patienten auch nach visueller Auswertung kein Skelettbefall vor und die Sensitivität des Programmes auf Patientenebene würde auf 98,5 Prozent ansteigen.

2. Auf Läsionsebene wurden 76,3 Prozent der Läsionen richtig erkannt. Bedeutsam erscheint in Bezug auf die Quantifizierung des Skelettbefalls, dass 40,2 Prozent der nicht vom Programm detektierten, aber in der visuellen Referenzbegutachtung als skelettale Hodgkinmanifestationen bewerteten Läsionen eine Größe von weniger als 0,25 ml besaßen. Der geringe prozentuale Anteil dieser sehr kleinen Läsionen von jeweils weniger als 1 Prozent an der gesamten skelettalen Tumorlast von im Mittel 22,30 ml erscheint für die Quantifizierung des Skelettbefalls daher vernachlässigbar. Außerdem kann man davon ausgehen, dass derart kleine Läsionen teilweise Artefakten entsprechen. Darüber hinaus konnten (Georgi et al. 2018) zeigen, dass derart kleine Läsionen unter Chemotherapie zuverlässig rasch verschwinden und somit bereits bei einer geringen Therapieintensität keine Rolle für den Erkrankungsverlauf spielen.
3. Bei der experimentellen Testung zur Bestimmung des optimalen Schwellenwertes an elf Patienten mit mindestens 5 Skelettläsionen (siehe Ergebnisse 6.2.5) wurde bei Anwendung von zweieinhalb Standardabweichungen über der durchschnittlichen Aktivitätsbelegung des Wirbelkörpers eine Sensitivität von 74,3 Prozent erreicht, die vergleichbar zur Anwendung bei 142 Patienten (76,3 Prozent) ist. Interessant ist hierbei, dass zwar nur 52,9 Prozent der flächig-diffusen, aber 92,6 Prozent der kräftig-fokalen Läsionen (siehe Tabelle 20 in Kapitel 6.2.5) erkannt wurden, wobei kräftige

multifokale Herdbefunde für das Hodgkin-Lymphom typisch sind (siehe Einleitung Kapitel 3.9 und (Purz et al. 2011)).

Auch (Muzahir et al. 2012) untersuchten 122 Patienten (Alter 6 bis 78 Jahre) mittels PET/CT und bilateraler Knochenmarkbiopsie an den Beckenkämmen, von denen 37 Patienten eine Mehrbelegung des Skeletts im PET aufwiesen, welche von (Muzahir et al. 2012) als Skelettbefall gewertet wurde. Sie unterteilten diese 37 Patienten nach Muster und Anzahl der FDG-Mehrbelegung in drei Gruppen:

1. homogene FDG-Mehrbelegung des Stammskeletts und proximaler Extremitätenknochen bei elf Patienten
2. unifokale Läsionen (ein oder zwei Skelettfoki mit FDG-Mehrbelegung) bei 6 Patienten)
3. multifokale Läsionen (mindestens drei Skelettfoki mit FDG-Mehrbelegung) bei 20 Patienten

Interessant ist, dass bei nur zwei der elf Patienten (18 Prozent) aus der ersten Gruppe eine Knochenmarkbeteiligung durch das Hodgkin-Lymphom mittels bilateraler Knochenmarkbiopsie nachgewiesen wurde, obwohl man ja bei Vorliegen eines Skelettbefalls und diffuser Aktivitätsmehrbelegung eine positive Biopsie an den zum Stammskelett gehörenden Beckenkämmen erwartet hätte. Dahingegen konnte in Gruppe 3 mit dem für das Hodgkin-Lymphom typischen Befallsmuster bei acht der zwanzig Patienten (40 Prozent) trotz einer möglichen Nichtbeteiligung der Beckenkämme ein höherer Anteil positiver Knochenmarkbiopsien erzielt werden als in Gruppe 1.

Daher liegt der Schluss nahe, dass zumindest bei einem Teil der in Gruppe 1 als Knochenmarkbefall bewerteten Patienten tatsächlich nur eine Knochenmarkaktivierung, die sich nach Therapie zurückentwickelte. Daher unterstützen auch die Ergebnisse von (Muzahir et al. 2012) die Bedeutung des für das Hodgkin-Lymphom typischen, kräftig multifokalen Befallsmusters.

Die pathologische Wertigkeit, der bei der visuellen Bewertung dokumentierten flächig-diffusen Befallsregionen ist nicht belegt und von der Kontrastaussteuerung des PET abhängig. Da eine klare Definition fehlt, ist der

Übergang zur Knochenmarkaktivierung fließend. Möglicherweise sind diese Regionen bei der visuellen Auswertung falsch-positiv bewertet worden. Betrachtet man diese Aktivitätsmehrbelegungen sehr sensitiv, wie (Muzahir et al. 2012) bei den elf Patienten in Gruppe A, so werden diese als Hodgkinmanifestationen bewertet.

(Kurch et al. 2013b) zeigte aber, dass bei Abgleich der homogenen Mehrbelegung des Stammskeletts im PET mit dem histologischen Ergebnis der Knochenmarkbiopsie kein Hodgkinbefall, sondern eine Knochenmarkaktivierung vorliegt.

Immerhin wurde ja aber die Hälfte dieser Läsionen auch vom Algorithmus erkannt. Das Programm kann zum einen sicher die kräftig-fokalen Läsionen mit hoher Spezifität aufzeigen. Darüber hinaus kann es als Interpretationshilfe zur Einordnung der flächig-diffusen Läsionen dienen und den Befunder in der Entscheidung unterstützen diese als Hodgkinläsionen oder Knochenmarkaktivierung zu bewerten.

4. Bei der kleinen prospektiv getesteten Kohorte mit Skelettbefall (11/142 Patienten) wurden die detektierten Läsionen erst nach Anwendung des Programmes mit den Ergebnissen der visuellen Referenzmethode verglichen. Hier lag die Sensitivität auf Läsionsebene sogar bei 90,2 Prozent (65/72 Läsionen wurden richtig angezeigt), was dem gefundenen Algorithmus hohe Detektionsgenauigkeit und Robustheit bescheinigt.

### **7.3.2 Verteilung der Skelettläsionen und Umfang der CT-Skelettmaske**

(Even-Sapir 2005) legen dar, dass über 90 Prozent der Skelettbefälle verschiedener Tumorleiden im roten Knochenmark auftreten. Dieses ist bei Erwachsenen nur noch im Stammskelett zu finden, bei Kindern und Jugendlichen – abhängig vom Alter – auch in peripheren Skelettanteilen. Die in dieser Arbeit untersuchten pädiatrischen Patienten zeigten einen deutlich bevorzugten Befall des Stammskeletts, wobei 90,5 Prozent der Läsionen auf Wirbelsäule, Skapula und Becken fielen, was konkordant zu den Ergebnissen von (Even-Sapir 2005) ist.

Die verwendete CT-Skelettmaske besteht aus dem Stammskelett und proximalen Extremitätenknochen. Diese Normierung erscheint zulässig, da beim Hodgkin-Lymphom nur selten ein Skelettbefall distal von Humerus- und Femurschaftmitte oder in der Schädelkalotte auftritt, sondern überwiegend das Stammskelett betrifft (Weiler-Sagie et al. 2014).

Zudem wird nach den EANM-Leitlinien bei Malignomen in der Regel der Ganzkörperscan vom Meatus acusticus externus bei elevierten Armen bis zur Femurschaftmitte durchgeführt (Boellaard et al. 2010), es sei denn ein Befall von Schädelkalotte oder distalen Extremitätenknochen ist wahrscheinlich.

(Weiler-Sagie et al. 2014) schlüsselten die Hodgkinläsionen bei 73 Patienten mit Knochenmarkbefall nach ihrer Lokalisation auf. 88 Prozent der Hodgkinmanifestationen fanden sich in Wirbelsäule und Becken. In der vorliegenden Arbeit waren diese beiden Regionen in 84 Prozent der Fälle ebenfalls am häufigsten befallen (siehe Ergebnisteil Kapitel 6.3 Unterpunkt 5).

Bei lediglich drei Patienten ließen sich Läsionen distal des oberen Humerusdrittels bzw. der Femurschaftmitte detektieren, wobei die distalen Extremitäten aufgrund des ausgedehnten Skelettbefalls im PET/CT-Scan mit erfasst worden waren. Bei diesen Patienten wurden daher auch die distalen Extremitätenknochen für die Läsionslast mit berücksichtigt, indem diese Läsionen mit den Algorithmus-Parametern nachvermessen wurden (vgl. Methodik 5.3 Unterpunkt 4 a). Da bei diesen drei Patienten zusätzlich mindestens drei Herde im Stammskelett vorhanden gewesen sind, hätte es zur Diagnosestellung eines Skelettbefalls nicht zwingend der Berücksichtigung der Läsionen in distalen Skelettabschnitten bedurft, vgl. (Purz et al. 2011; Kluge und Körholz 2011).

Auf der anderen Seite wurde durch die zusätzliche Bestimmung des Volumens dieser Läsionen eine korrekte Quantifizierung des Skelettbefalls gewährleistet.

### **7.3.3 Einschränkungen und Verbesserungsmöglichkeiten**

Der Algorithmus, der in dieser Arbeit entwickelt wurde, ist verschiedenen Einschränkungen unterworfen.

## 1. Nicht in die CT-Skelettmaske eingeschlossene skelettale Läsionen

Anteile des Skeletts wurden nicht in die CT-Skelettmaske integriert, wenn diese den HU-Schwellenwert unterschritten. Das betraf besonders die Spongiosa langer Röhrenknochen. Da durch gelbes Fettmark ersetzte Knochenmarkhöhlen in Femur und Humerus teils sogar negative Hounsfielddichtewerte annehmen, wurden diese nicht in der CT-Skelettmaske erfasst. In unserer Arbeit lagen dadurch 15 skelettale Läsionen in Humeri und Femora außerhalb der CT-Skelettmaske (entspricht 6,2 Prozent aller 241 nicht detektierten Läsionen). Kleinere Fehlstellen z. B. in den Wirbelkörpern wurden durch Anwendung der „Hole Filling Filter“-Funktion (vgl. Kapitel 6.1.2 und 7.1) in die CT-Skelettmaske integriert. Allerdings ist die „Hole Filling Filter“-Funktion nicht in der Lage, die Knochenmarkhöhlen der großen Röhrenknochen in die CT-Skelettmaske einzuschließen, da die in diesen Markhöhlen befindlichen Voxel nicht von entsprechend vielen, bereits zur CT-Skelettmaske gehörigen Voxeln umgeben sind, um selbst als Skelettvoxel erkannt zu werden (siehe Abbildung 27 im Kapitel 6.1.2).

## 2. Differenzierbarkeit von hämatogenem Skelettbefall gegenüber E-Befall

Anders als bei hämatogenem Befall, bei dem sich der Tumor stets vom Knochenmark beginnend ausbreitet und die Kortikalis nur in einem Teil der Fälle mit betrifft, wächst bei E-Befall der Tumor von außen in den Knochen hinein und betrifft somit immer auch die Kortikalis. Sechs von 142 Patienten (4,2 Prozent) wiesen Skelettbefall ausschließlich als E-Befall durch Wachstum per continuitatem auf. E-Befall kann momentan durch den Algorithmus nicht von hämatogenem Skelettbefall differenziert werden. Selbst erfahrenen Untersuchern fällt diese Differenzierung gelegentlich schwer. Da der größere Anteil der Tumormasse und damit auch der Aktivitätsbelegung bei E-Befall außerhalb der CT-Skelettmaske liegt, könnte eine Funktion etabliert werden, welche bestimmen kann, ob der größere Anteil der Mehrbelegung innerhalb oder außerhalb der CT-Skelettmaske liegt. Bei überwiegend extraskelletaler Mehrbelegung würde ein E-Befall vom Programm ausgegeben werden. Dadurch könnte eine gesonderte Erfassung dieses selteneren Skelettbefalls gelingen, der zum einen eine deutlich bessere Prognose hat und nicht zur Einordnung der Erkrankung ins Stadium IV nach Ann-Arbor führt (siehe Einleitung 3.6). Eine Kortikalisbeteiligung

kann von therapeutischer Relevanz sein, da im Falle einer Strahlentherapie nur ossäre Befälle, nicht aber Knochenmarksbefälle mit bestrahlt werden (Kurch et al. 2013b).

### 3. Sensitivität des Programmes in Bezug auf problematische Skelettregionen

Der Algorithmus zeigte unterschiedliche Sensitivitäten in Abhängigkeit von der betrachteten Skelettregion. Läsionen in den Rippen und den Scapulae wurden hierbei mit deutlich geringerer Sensitivität im Vergleich zu den übrigen Regionen detektiert: Der Anteil der falsch-negativen Läsionen in Bezug auf die Gesamtzahl der Läsionen in diesen beiden Regionen beträgt mehr als 60 Prozent (vgl. Ergebnisse, Kapitel 6.3 d, Tab. 28), was einer Sensitivität in diesen beiden Skelettregionen von weniger als 40 Prozent entspricht.

Das lag zum einen an der geringeren Größe der Läsionen im Vergleich zu anderen Skelettregionen. Besonders die Läsionen in den Rippen überschritten das Grenzvolumen von 0,25 ml nicht. Zudem gilt, dass kleine Läsionen deutlich anfälliger für Partialvolumeneffekte (PVE) und die daraus resultierende Unterschätzung der Aktivitätsbelegung sind (Soret et al. 2007). Das führt dazu, dass kleine Läsionen den ermittelten Grenzwert der Aktivitätsbelegung ( $SUV_{\text{mean}} + 2,5 \text{ SD}$  des Referenzwirbelkörpers) nicht überschreiten und dadurch nicht detektiert werden.

Außerdem liegt ein möglicher weiterer Fehler in der Anfälligkeit des Thoraxskeletts für Bewegungsartefakte durch die Atemexkursionen des Patienten (Nehmeh und Erdi 2008). CT-Aufnahmen können aufgrund der sehr kurzen Untersuchungszeit von wenigen Sekunden in Atemanhaltetechnik angefertigt werden, während PET-Aufnahmen aufgrund der längeren Dauer bei normalem Atemzyklus aufgenommen werden (Rodríguez-Vigil et al. 2006). Dadurch sind besonders die Rippen und das Sternum anfällig für Atembewegungen, wodurch die Aktivitätsbelegung aus der PET-Komponente dieser Regionen zum Teil außerhalb der CT-Skelettmaske zu liegen kommt und deshalb nicht vollständig vom Algorithmus erfasst werden kann.

#### 4. Zusätzlich detektierte Aktivitätsmehrbelegungen

##### a) Extraskelettale Aktivitätsmehrbelegungen

Da bei 62 Patienten (43,7 Prozent von 142 Patienten) Blasenaktivität auf die CT-Skelettmaske projiziert wurde und bei 72 Patienten (51,7 Prozent von 142 Patienten) Spillover der Aktivität von Herz, braunem Fett oder nodalem Tumorgewebe durch den Algorithmus als falsch-positive Läsion detektiert wurde, entsteht ein relevanter Mehraufwand für den Programmanwender, weil er diese Läsionen manuell ablehnen muss. Daher wäre es von besonderer Bedeutung, eine Funktion im Programm zu implementieren, welche Gewebe mit hoher Aktivitätskonzentration, deren Volumen überwiegend außerhalb der CT-Skelettmaske lokalisiert ist, nicht mehr in diese integriert. Diese Funktion würde gleichzeitig auch zur Unterscheidung zwischen E-Befall und hämatogenen Skelettbefall dienen, wie unter 7.3.3, Unterpunkt 1 skizziert. Dadurch würde die Bearbeitungszeit verkürzt und die Praktikabilität für die klinische Routine gesteigert werden.

Da Kinder und Jugendliche deutlich mehr braunes Fettgewebe zur Wärmeerzeugung als Erwachsenen besitzen, kommt es in 15 bis 20 Prozent aller PET-Untersuchungen bei Kindern und Jugendlichen zu einer Mehranreicherung von  $^{18}\text{F}$ -FDG in aktiviertem braunem Fettgewebe (Stauss et al. 2008). Die Aktivierung von braunem Fettgewebe erfolgt durch Kälte und einen erhöhten Sympathikotonus (Boellaard et al. 2008). Daher sollte besonderer Wert auf Maßnahmen in der Patientenvorbereitung gelegt werden, welche die Aktivierung braunen Fettgewebes präventiv unterdrücken wie Wärmeapplikation oder unselektive Beta-Blocker, siehe (DGN-Geschäftsstelle).

##### b) Skelettale Aktivitätsmehrbelegungen

Wie im Methodikteil 5.3 im Unterpunkt 5d erläutert, wurden alle Datensätze mit Auswertung der skelettalen Aktivitätsmehrbelegungen gespeichert. Von besonderem Interesse erscheint das für eine Zweitbegutachtung der vom Algorithmus gegenüber der visuellen Referenzbegutachtung zusätzlich detektierten skelettalen Aktivitätsmehrbelegungen. Möglicherweise handelt es sich um flächig-diffuse Mehranreicherungen, die auch in der visuellen Referenzbegutachtung als zumindest grenzwertiger Skelettbefall beurteilt

werden würden. Andernfalls müssten diese Mehrbelegungen als falsch-positive Skelettläsionen eingestuft werden. Die gezielte Reevaluation dieser zusätzlich erkannten Läsionen durch einen Facharzt für Nuklearmedizin ist Bestandteil eines Anschlussprojektes. Das Programm kann in jedem Fall eine Detektionshilfe für den befundenden Nuklearmediziner sein, indem es ihm eine Vorauswahl von Skelettläsionen anbietet, die dieser akzeptieren oder verwerfen kann. Die Verwendung des Programmes kann somit zusätzliche Sicherheit in der klinischen Routine bringen, vgl. (Liang et al. 2016) beim Lungenkrebsscreening im CT.

#### **7.4 Bedeutung (semi-)automatischer Algorithmen in der funktionellen Bildgebung und Relevanz von hämatogenem Skelettbefall beim Hodgkin-Lymphom und soliden Tumoren**

Wie bereits beschrieben, ist bisher kein Programm vorhanden, welches hämatogenen Skelettbefall in einer PET/CT-Untersuchung im gesamten Skelett halbautomatisch detektieren kann und insbesondere keines, welches an pädiatrischen Patienten mit Hodgkin-Lymphom getestet und angewendet wurde. Dass ein solches bisher nicht entwickelt wurde, mag auch daran liegen, dass zwar das Vorhandensein (Einordnung ins Stadium 4 nach der Ann-Arbor-Klassifikation bei hämatogenem Befall), aber nicht das Ausmaß des Skelettbefalls als relevanter Prognosefaktor beim Hodgkin-Lymphom ermittelt wurden.

Hingegen wurde für solide Tumoren gezeigt, dass das Ausmaß einer Skelettmetastasierung wesentlichen Einfluss auf die Prognose hat. Patienten mit Mamma- oder Prostatakarzinom zeigen besonders häufig osteoblastische Fernmetastasen. Bei diesen Patienten konnte gezeigt werden, dass die mittels Skelettszintigrafie bestimmte skelettale Metastasenlast mit dem Überleben korreliert. Hierzu wurde bereits in der Einleitung in Kapitel 3.7 auf eine von der Firma EXINI Diagnostics entwickelte Software eingegangen, welche es ermöglicht, eine Atlas-basierte Segmentierung von Skelettszintigrafiebildern vorzunehmen (Sjostrand et al. 2009) und anschließend den Skelettbefall zu quantifizieren.

Dieses Programm hat somit einen ähnlichen Zweck wie das in der vorliegenden Arbeit verwendete, nämlich die Prognoseabschätzung des quantifizierten Skelettbefalls. Da auf Lymphome angewandte, dreidimensionale Programme zur halbautomatischen



Auswertung des Skelettbefalls fehlen, sollen folgend die bisherigen Erkenntnisse beim Mamma- und Prostatakarzinom wie auch das verwendete Programm der Firma EXINI Diagnostics vorgestellt werden. Ziel ist, trotz prinzipieller Unterschiede, das hier entwickelte Programm mit dem bei Skelettszintigrafiebildern angewandten Programm zu vergleichen.

Zu den prinzipiellen Unterschieden gehören:

- 1) Der dargestellte Stoffwechselprozess (Glukosestoffwechsel bei der  $^{18}\text{F}$ -FDG-PET versus Ostoblastenstoffwechsel bei der Skelettszintigrafie)
- 2) Die Auswertung von 3D-Datensätzen bei der  $^{18}\text{F}$ -FDG-PET versus planaren, 2D-Datensätzen bei der Skelettszintigrafie
- 3) Die Krankheitsentität (Prostata-/Mammakarzinome als solide Tumore versus Lymphom als generalisierte hämatologische Erkrankung)
4. Unterschiede hinsichtlich der Patientengruppe (Kinder mit deutlich weniger degenerativen Skelettveränderungen vs. ältere Erwachsene mit relativ häufigen degenerativen Skeletterkrankungen)

#### **7.4.1 Relevanz des mittels Skelettszintigrafie bestimmten Skelettbefalls beim Prostatakarzinom**

(Soloway et al. 1988) stellten beim Prostatakarzinom eine visuelle Skala zur Abschätzung der Ausprägung einer metastatischen Skeletterkrankung bei Durchführung einer Skelettszintigrafie vor. Sie teilten ihre Patienten nach der Anzahl der Skelettmetastasen in fünf Schweregrade ein. Diese wurden bei 166 Männern mit skelettal metastasiertem Prostatakarzinom angewendet. Die fünf Kategorien sind in nachfolgender Tabelle 31 aufgeführt und dem jeweiligen 2-Jahres-Überleben gegenübergestellt.

<b>Soloway</b>	<b>Skelettmetastasen</b>	<b>2-Jahres-Überleben</b>
0	Keine	100%
1	< 6	94%
2	6-20	74%
3	> 20	68%
4	Super Scan	40%

Tab. 31: Stratifikation der Patienten mit Skelettmetastasen bei Prostatakarzinom nach (Soloway et al. 1988)

(Soloway et al. 1988) setzten die Grenze zwischen Gruppe 1 und 2 bei 6 Skelettmetastasen und zwischen Gruppe 2 und 3 bei 20 Skelettmetastasen. Sehr große Skelettmetastasen, deren Größe das Volumen von 50 Prozent eines Wirbelkörpers überschreiten, werden als zwei Läsionen gewertet. Einen Sonderfall und die höchste Stufe dieser Einteilung stellt die Soloway-Gruppe 4 bei Vorliegen eines „Superscans“ dar. Dieser entspricht einer deutlich verminderten bis fehlenden Aktivitätsbelegung in den Weichteilen, Nieren und ableitenden Harnwegen aufgrund einer disseminierten Skeletterkrankung.

Das 2-Jahres-Überleben beträgt für Patienten ohne Skelettmetastasen 100 Prozent und nimmt mit zunehmender Metastasenzahl ab (siehe oben stehende Tabelle 31). (Soloway et al. 1988) konnten für die Gruppen Soloway 1 (weniger als sechs Skelettmetastasen) und Soloway 4 (Superscan) nachweisen, dass sich diese im 2-Jahres-Überleben deutlich voneinander und von den übrigen Gruppen unterschieden. Allerdings zeigte sich kein signifikanter Unterschied im 2-Jahres-Überleben zwischen den Gruppen Soloway 2 (sechs bis 20 Skelettmetastasen) und Soloway 3 (mehr als 20 Skelettmetastasen). Als Referenzmaß für eine Läsion gilt die halbe Fläche eines Wirbelkörpers, ist eine Mehranreicherung größer, wird sie wie zwei Läsionen gewertet. Da die Skelettläsionen gezählt und in ihrer Größe geschätzt werden müssen, ist dieses Verfahren bei multiplen Metastasen zeitaufwendig und bei großen Läsionen z. B. im Becken auch ungenau. Problematisch ist diese Stratifizierung auch bei Verlaufsuntersuchungen, wo Läsionen konfluieren können und dadurch die Zählbarkeit erschwert wird.

Beim Hodgkin-Lymphom hingegen zeigt sich typischerweise ein multifokales Befallsmuster und die Skelettläsionen konfluieren nicht (Purz et al. 2011). Darüber hinaus verschwinden beim üblicherweise guten Therapieansprechen auf die Chemotherapie die meisten Läsionen bereits in der Frühresponsebeurteilung im ERA-PET (Georgi et al. 2018) und sind dadurch visuell leichter zählbar.

Zur Quantifizierung des Skelettbefalls in der Skelettszintigrafie, insbesondere beim skelettal metastasierten Prostatakarzinom, hat sich der Bone Scan Index (BSI) etabliert. Dieser wurde von (Imbriaco et al. 1998) bei 90 Patienten und 263 Skelettszintigrafien mit fortgeschrittenem, androgenresistentem Prostatakarzinom erstmals angewandt. Hierbei wird das Ausmaß des Befalls jedes Knochens visuell abgeschätzt, um anschließend durch Summation aller Herde den Skelettbefall zu

bestimmen, wobei der prozentuale Anteil der einzelnen Knochen am Gesamtskelett bekannt ist. Man bezieht sich hierbei auf Standardtabellen, die die International Commission on Radiological Protection (ICRP) veröffentlicht hat (vgl. Publication 23: Report of the task group on reference man, Download möglich unter <http://www.icrp.org/publication.asp?id=ICRP%20Publication%2023>, zuletzt geprüft am 15.11.2018).

(Imbriaco et al. 1998) konnte akzeptable Werte sowohl für die Intra- als auch die Interobservervariabilität bestimmen, letztere lag bei 15 Prozent. Problematisch ist bei manueller Erhebung und je nach Anzahl der Läsionen die erforderliche Zeit von 5 bis 30 Minuten, welche die Ressourcen in der klinischen Routine überfordert.

(Erdi et al. 1997) stellten ein auf der BSI-Methode basierendes, halbautomatisches Analyse-Programm für Skelettmetastasen („BMets analysis“) vor, welches die semiautomatische Analyse planarer Skelettszintigrafien mit anteriorer und posteriorer Projektion erlaubt.

Hierzu wird zunächst das Volumen des gesamten Skeletts anhand von Alter, Größe, Geschlecht und Rasse mittels linearer Regression geschätzt. Anschließend wählt man die einzelnen Läsionen in der Skelettszintigrafie im Programm mittels Mausklick aus. Ein „RegionGrow“-Algorithmus fügt dann zur Skelettläsion gehörige Nachbapixel bis zum Erreichen eines errechneten, kontrastabhängigen, optimalen Schwellenwertes hinzu. Dieser kontrastabhängige Schwellenwert wird anhand von initialen Phantom-Studien mit vorgegebenen Tracerkonzentrationen gegenüber einem Hintergrund ermittelt. Dann wird der Anteil der so vermessenen Skelettläsion am jeweiligen Knochen bestimmt. Anhand von Standardtabellen der International Commission on Radiological Protection (siehe oben) wird der Anteil des befallenen Knochens am Gesamtskelett ermittelt. Anschließend verglichen (Erdi et al. 1997) unter Analyse von zehn Skelettszintigrafien von sechs Patienten die beschriebene, halbautomatische Methode mit der visuellen BSI-Methode. Sie konnten hierbei zeigen, dass der Anwender bei der visuellen Methode (vgl. (Imbriaco et al. 1998)) systematisch die Läsionsgröße überschätzt. Eine Unterschätzung durch das Programm konnte man anhand der vorangegangenen Eichung auf die untersuchten Phantome mit bekannter Größe und Aktivität ausschließen. Limitationen des Region-Grow-Algorithmus sind bei einem geringen Signal-zu-Rausch-Verhältnis gegeben, da dann die Grenze der Skelettläsion nicht mehr eindeutig durch den „RegionGrow“-Algorithmus erkannt wird.

Außerdem ist es nicht möglich, einzelne Schädelknochen voneinander zu trennen, da diese vom verwendeten Programm als ein einziger Knochen erkannt werden. Die Dauer dieser halbautomatischen Methode beträgt 1 bis 4 Minuten pro Patienten, abhängig von Anzahl und Erkennbarkeit der Läsionen. Damit gelingt (Erdi et al. 1997) eine deutliche Zeitersparnis gegenüber der manuellen BSI-Methode. Interobservervariabilität tritt nur noch auf, wenn Uneinigkeit über das Vorhandensein des Befalls eines Knochens besteht, sodass dieser per Mausklick angewählt wird oder nicht. Ein besonderer Vorteil der reduzierten Interobservervariabilität besteht in der objektiven Auswertung von Verlaufsuntersuchungen zur Evaluation des Therapieansprechens.

Die Firma EXINI Diagnostics hat eine Software entwickelt, welche es ermöglicht, eine Atlas-basierte Segmentierung von Skelettszintigrafiebildern vorzunehmen (vgl. Einleitung 3.7). Anschließend detektiert das Programm Mehranreicherungen als Läsionen und es kann nach dem Prinzip von (Erdi et al. 1997) eine automatische, der BSI-Methode entsprechende Quantifizierung des Skelettbefalls beim Prostata- und Mammakarzinom vorgenommen werden (siehe <http://bonescanindex.org/about-bsi/>, Stand 15.11.2018).

(Ulmert et al. 2012) wendeten das EXINI-Programm als erster an und stellt die sehr kurze Anwendungszeit der automatischen Methode von 10 Sekunden gegenüber 5 bis 30 Minuten bei Anwendung der manuellen BSI-Methode heraus. Bei 133 Patienten konnte er eine gute Übereinstimmung zwischen manueller und automatischer Methode zeigen. Neben PSA, Gleason-Score und klinischem Stadium der Erkrankung konnte der automatisierte BSI als unabhängiger Prädiktor für das Überleben des ossär metastasierten Prostatakarzinoms nachgewiesen werden. Allerdings wurde das Ausmaß der Skeletterkrankung bei einem sehr hohen BSI ( $>10$ ) und somit ausgeprägten Befall durch das EXINI-Programm im Gegensatz zur manuellen Methode unterschätzt.

In der zweiten Version des Programmes konnte durch Optimierung der Programmparameter die exakte Erfassung auch bei sehr fortgeschrittener skelettaler Metastasierung des Prostatakarzinoms nachgewiesen werden (Anand et al. 2016a). Auch bei Patienten mit metastasiertem Prostatakarzinom unter Enzalutamidtherapie konnte ein zusätzlicher Nutzen des automatisierten BSI zusätzlich zum etablierten

prostataspezifischen Antigen als Prädiktor für das Gesamtüberleben gezeigt werden (Anand et al. 2016b).

(Larson 2016) postulieren, dass das EXINI-Programm beim knöchern metastasierten Prostatakarzinom aktuell von besonderem Nutzen ist, da verschiedene neue Medikamente zur palliativen Lebenszeitverlängerung vorhanden sind und ein erweitertes Monitoring vonnöten sein wird. Die Autoren betonen außerdem, dass dieser computerbasierte Algorithmus prinzipiell auch auf andere an den Knochen anheftende Radiopharmaka übertragen werden kann.

#### **7.4.2 Relevanz des mittels PET/CT bestimmten Skelettbefalls beim Hodgkin-Lymphom**

Mit dem hier vorgestellten Algorithmus ist ein Werkzeug entwickelt worden, welches es erstmals erlaubt, das metabolische Ausmaß des Skelettbefalls bei pädiatrischen Patienten mit Hodgkin-Lymphom im PET/CT zu quantifizieren. Hierzu wurde das ermittelte Volumen der Skelettläsionen auf das Gesamtvolumen der normierten Skelettmaske bezogen und damit ein Wert analog zum BSI ermittelt.

Danach wurde in der hier vorliegenden Arbeit das ereignisfreie Überleben von 101 Patienten in Abhängigkeit vom quantifizierten Skelettbefall nach 36 Monaten betrachtet, wobei alle Patienten mindestens eine Skelettläsion aufwiesen. Es konnte allerdings keine Relevanz des mittels PET/CT quantifizierten Skelettbefalls für das ereignisfreie Überleben gezeigt werden (siehe Ergebnisse 6.4).

Das kann durch die stadienabhängige, risikoadaptierte Therapie des Hodgkin-Lymphoms erklärt werden, wobei Skelettbefall mit der Einordnung ins Stadium IV der Ann-Arbor-Klassifikation eine intensivierte Chemotherapie nach sich zieht (Olson und Donaldson 2008). Da die verabreichte Chemotherapie durch die gute Durchblutung des roten Knochenmarks zügig zu den chemosensitiven Lymphomabsiedlungen gelangt, sind die meisten Skelettläsionen bereits im ERA-PET nicht mehr nachweisbar, vgl. (Georgi et al. 2018). Bei Knochenbefall werden die Kortikalisherde zusätzlich bestrahlt, siehe (Kurch et al. 2013b).

Es gibt in der einschlägigen Literatur nur wenige Publikationen, welche Skelettbefall und seinen Einfluss auf das ereignisfreie und Gesamtüberleben speziell bei pädiatrischen Patienten mit Hodgkin-Lymphom untersucht haben:

(Hines-Thomas et al. 2010) analysierten eine Kohorte von 228 pädiatrischen Patienten. Bei 21 Patienten wurde durch eine Knochenmarkbiopsie am Beckenkamm ein Skelettbefall diagnostiziert, das von (Purz et al. 2011) mittels PET/CT dargestellte, typische multifokale Befallsmuster des Hodgkin-Lymphoms wurde nicht berücksichtigt. Sie konnten keine signifikante Korrelation zwischen Knochenmarkbefall und Gesamtüberleben oder ereignisfreiem Überleben nachweisen. Bei 3,1 Prozent der Patienten wurde die Stadieneinteilung durch die positive Knochenmarkstanze zwar geändert, es wurden aber keine Auswirkungen auf risikoadaptierte Therapie oder Prognose beobachtet.

Das Hodgkin-Lymphom mit seinem multifokalen Befallsmuster spart häufig den Beckenkamm aus. Die geringe Sensitivität der Knochenmarkstanze an den Beckenkämmen könnte die Ergebnisse von (Hines-Thomas et al. 2010) entscheidend beeinflusst haben. Im Gegensatz dazu wurden in der vorliegenden Arbeit **alle** mittels **PET/CT** erfassten **Skelettläsionen** berücksichtigt. Bei Vergleich von Gesamt- und ereignisfreiem Überleben der pädiatrischen Patienten zeigte sich allerdings auch hier kein signifikanter Unterschied zwischen den verschiedenen Ausprägungen des Skelettbefalls.

Einige wenige Publikationen untersuchten erwachsene Patienten mit Hodgkin-Erkrankung und Skelettbefall. So stellten (Feltl et al. 2008) die Ergebnisse bei 198 erwachsenen Patienten mit Hodgkin-Lymphom vor. Bei 14 Patienten war ein hämatogener Skelettbefall nachweisbar. Dieser wurde in zwei Fällen nur mittels Szintigrafie, in acht Fällen nur mittels CT, in drei Fällen mittels CT und MRT und in nur einem Fall mittels CT und PET diagnostiziert. Eine routinemäßige Knochenmarkstanze wurde nicht zur Diagnosestellung eingesetzt. Das ereignisfreie Überleben im Zwei-Jahres-Zeitraum war mit 71,4 Prozent der Patienten mit Skelettbefall signifikant verringert gegenüber 92,7 Prozent bei Patienten ohne Skelettbefall. Das Gesamtüberleben im Zwei-Jahres-Zeitraum war wie in der vorliegenden Arbeit unverändert. Da (Feltl et al. 2008) weder eine MRT- noch eine PET-Untersuchung routinemäßig durchführte, wurde im Gegensatz zur EuroNet-PHL-C1-Studie und der vorliegenden Arbeit ein früher Knochenmarkbefall wahrscheinlich gar nicht erfasst.

(Hasenclever et al. 1998) zeigten bei erwachsenen Patienten mit Hodgkin-Lymphom, dass Knochenmarkbefall häufig auch mit ungünstigen Prognosefaktoren wie Stadium IV in anderen Organsystemen, großer Tumorlast und B-Symptomen assoziiert ist. Das Ausmaß des Skelettbefalls spielte dann keine zusätzliche Rolle für die Prognose.

Die Therapie des pädiatrischen Hodgkin-Lymphoms wurde in den letzten Jahren im Rahmen von Therapieoptimierungsstudien sukzessive verbessert (siehe Einleitung 3.3). Skelettbefall führt zur Zuordnung eines höheren Krankheitsstadiums und einer Therapieintensivierung, wodurch alle Skelettläsionen trotz fortgeschrittener Erkrankung ein Therapieansprechen (Georgi et al. 2018) zeigen. Daher hat die hier vorliegende Arbeit einen besonderen Fokus darauf gelegt, ob es innerhalb der Gruppe mit Skelettbefall Patienten gibt, die aufgrund einer sehr ausgeprägten skelettalen, initialen Tumormasse ein schlechteres Outcome zeigen. Allerdings wurden weder das ereignisfreie noch das Gesamtüberleben durch das Ausmaß des Skelettbefalls signifikant beeinflusst.

#### **7.4.3 Ansätze zur (semi-)automatischen Quantifizierung der Ganzkörper tumorlast bei Lymphomen mittels PET/CT**

In den medizinischen Literaturdatenbanken wurde keine Arbeit gefunden, welche den prognostischen Wert des mittels PET/CT ermittelten metabolischen Lymphomvolumens im Skelett analysiert hat. Allerdings gibt es Arbeiten, die versuchen, die Ganzkörper tumorlast bei Lymphomen im Initialstaging zu quantifizieren:

(Meignan et al. 2015) zeigen in ihrem Review verschiedener Studien zwei Methoden auf, das metabolische Gesamttumorvolumen bei Patienten im Initialstaging zu ermitteln.

Einige Arbeitsgruppen verwenden einen absoluten Grenzwert des SUV von 2,5. Alle Mehranreicherungen, die diesen Wert überschreiten, werden als Hodgkin-Befall gewertet. So untersuchten (Song et al. 2012) das **metabolische Tumervolumen (MTV)** bei 165 Patienten mit einem primär gastrointestinalen, diffusen, großzelligen B-Zell-Lymphom. Sie konnten zeigen, dass das MTV ein besserer Prognosefaktor für

ereignisfreies und Gesamtüberleben als der  $SUV_{max}$  ist. Dieser Ansatz überschätzt das Tumolvolumen jedoch erheblich, indem er Hintergrundaktivität mit einbezieht. Diese Methode erscheint besonders ungeeignet, um im Skelett angewendet zu werden, da hier die Aktivitätsbelegung allein durch Knochenmarkaktivierung deutlich gesteigert sein kann, wie auch die hier vorliegende Arbeit gezeigt hat.

Andere Arbeitsgruppen definieren den Grenzwert in Abhängigkeit vom  $SUV_{max}$ . (Kanoun et al. 2014) kalkulierten das metabolische Tumolvolumen in einer retrospektiven Single-Center-Studie bei 59 Hodgkin-Patienten im Alter von 16 bis 76 Jahren. Hierbei errechneten sie das MTV als Summe aller Hodgkinläsionen, die 41 Prozent des  $SUV_{max}$  der jeweiligen Läsion überschritten. Sie zeigten, dass das MTV ein besserer Prognosefaktor als der größte Durchmesser des mediastinalen Tumorbulks ist. Zudem zeigten sie, dass das initiale metabolische Tumolvolumen auch ein unabhängiger Risikofaktor zusätzlich zum Ansprechen im ERA-PET ist. Sie konnten ein metabolisches Tumolvolumen von 225 ml als optimalen Cut-Off ermitteln. Hierbei betrug das ereignisfreie Überleben 85 Prozent bei einem  $MTV < 225$  ml gegenüber 42 Prozent bei einem  $MTV \geq 225$  ml und das durch die Erkrankung bedingte Gesamtüberleben 88 Prozent bei einem  $MTV < 225$  ml gegenüber 45 Prozent bei einem  $MTV \geq 225$  ml. Eine homogene Mehrbelegung des Skeletts, welche möglicherweise einer Knochenmarkaktivierung entsprach, schlossen sie nicht als Läsionen ein. Diese Arbeitsgruppe hatte keine Strategie das skelettale Gesamttumolvolumen und seine Bedeutung für die Prognose zu ermitteln. Dies wird durch das Programm der vorliegenden Arbeit für das pädiatrische Hodgkin-Lymphom zur Verfügung gestellt.

Auch (Cottreau et al. 2016) kalkulierten das MTV anhand von 41 Prozent des  $SUV_{max}$  der jeweiligen Läsion und führten eine retrospektive Analyse bei 108 Patienten mit einem peripheren T-Zell-Lymphom durch. Sie zeigten, dass bei einem initialen  $MTV > 230$  ml das progressionsfreie Zwei-Jahres-Überleben mit 26 Prozent deutlich gegenüber 71 Prozent bei einem  $MTV < 230$  ml reduziert war. In einer multivariaten Analyse war das initiale, metabolische Tumolvolumen der einzige unabhängige Risikofaktor für das progressionsfreie und Gesamtüberleben. Auch diese Methode ist natürlich nicht in der Lage isoliert die Gesamtskeletttumorlast im Gegensatz zur vorliegenden Arbeit zu quantifizieren.



(Ceriani et al. 2015) untersuchten 103 Patienten mit primär mediastinalem großzelligem B-Zelllymphom im Initialstaging und konnten zeigen, dass das MTV in der univariaten Analyse ein unabhängiger Risikofaktor ist. Die Grenze der Läsionen wurde anhand von 25 Prozent des  $SUV_{max}$  definiert. In der multivariaten Analyse war das totale Glykolysevolumen als Produkt aus metabolischen Tumervolumen und  $SUV_{max}$  der stärkste prognostische Faktor. Probleme zeigte ihre Methode bei kleineren, nicht im Mediastinum liegenden Läsionen, z. B. peripheren Lymphknoten, deren Volumen aufgrund eines a. e. durch Partialvolumeneffekt erniedrigten  $SUV_{max}$  falsch erhöht war. Diese Methode erscheint daher insbesondere ungeeignet, kleinere skelettale Manifestationen zu detektieren.

Keiner dieser bisher angewandten Algorithmen kann den Skelettbefall in der PET/CT volumetrieren und damit die skelettale Gesamtumorlast bestimmen, um diese als Prognosefaktor zu testen. Genau das aber wurde in der vorliegenden Arbeit für das pädiatrische Hodgkin-Lymphom erprobt. Perspektivisch kann der hier vorgestellte Algorithmus auch für die **Skeletttumormassenvolumetrie** bei verschiedenen **NHL getestet** werden.

### **7.5 Weitere Anwendungsmöglichkeiten und mögliche Bedeutung des entwickelten Algorithmus in der klinischen Routine**

Die vorgestellte Methode ist zur Zeit die einzige, welche bei malignen Erkrankungen spezifisch die Erfassung von Skelettbefall mittels PET/CT ermöglicht.

Der entwickelte Algorithmus stellt ein wertvolles Werkzeug zur semiautomatischen Beurteilung von Skelettbefall beim PHL dar. Der bisher ausschließlich visuell ermittelte Skelettbefall kann quantifiziert und das erheblich schwankende Ausmaß des Skelettbefalls der Patienten besser verglichen werden.

Momentan ist bei der Detektion verschiedener skelettaler Lokalisationen bei Hodgkin-Lymphom noch keine Zeitersparnis gegenüber der visuellen Auswertung gegeben wie sie die EXINI-Software beim skelettal metastasierten Prostatakarzinom bietet. Bei

weiterer Optimierung des Programmablaufs kann aber zumindest eine deutliche Reduktion der Bearbeitungszeit gelingen.

Hierzu müsste insbesondere die Erstellung der Skelettmaske optimiert werden. Größere Fehlstellen von Skelettanteilen mit zu geringer Dichte wie Knochenmarkhöhlen müssten automatisch in die CT-Skelettmaske integriert werden, sodass diese nicht manuell nachvermessen werden müssen (vgl. Ergebnisse, Kapitel 6.3 Unterpunkt 5 d und Diskussion, Kapitel 7.3.3 Unterpunkt 1). Desweiteren sollte eine automatische, iterative Anwendung der Knochenfüllerfunktion (vgl. Ergebnisse Kapitel 6.1.2) erfolgen, um eine vollständige CT-Skelettmaske zu erhalten. Zum anderen sollten die metallische Patientenliege und orales Kontrastmittel automatisch vom Programm als extraskelletal erkannt werden.

Besonders hervorzuheben ist die nachgewiesene Eignung für den Einsatz in einer Multi-Center-Studie, da sonst in vielen Fällen die ermittelten Schwellenwerte auf die jeweilige untersuchende Einrichtung beschränkt bleiben (Thie 2004). Der in der vorliegenden Arbeit beschriebene Algorithmus wurde im Gegenzug dazu bereits auf zahlreiche Datensätze aus bis zu 200 verschiedenen Zentren der Multicenterstudie EuroNet-PHL-C1 angewandt und erscheint dadurch auch für eine Verwendung in der Folgestudie EuroNet-PHL-C2 geeignet. Da in dieser Studie keine Knochenmarkstanze zur Diagnose von Knochenmarkbefall mehr angewendet wird, kommt der korrekten Detektion des Knochenmarkbefalls in der PET eine besondere Bedeutung zu (Kluge und Körholz 2011). Hilfreich könnte dieser Algorithmus dann als Zweitbegutachtung neben der visuellen Beurteilung sein, da die Diagnose des Knochenmarkbefalls hauptsächlich auf der PET/CT-Untersuchung beruht.

Denkbar ist auch eine Testung und Anwendung des Algorithmus bei erwachsenen Patienten mit Hodgkin-Lymphom.

(El-Galaly et al. 2012) untersuchten bei 454 erwachsenen Patienten mit Hodgkin-Lymphom, welche auch ein PET/CT erhalten hatten, den Nutzen der Knochenmarkbiopsie in Hinblick auf das Staging. Der zusätzliche Nutzen der Knochenmarkstanze für die Diagnostik wurde hier als minimal eingeschätzt, da kein Patient aufgrund der Biopsie einer höheren Therapiegruppe zugewiesen wurde. Weiterhin erzielte keine Knochenmarkstanze bei Patienten im Stadium I oder II der Ann-Arbor-Klassifikation nach PET/CT einen positiven Befund. Die Autoren weisen

deshalb darauf hin, dass eine Nichtanwendung der Knochenmarkstanze bei keinem Patienten eine Änderung der Behandlungsstrategie zur Folge gehabt hätte.

Umso wichtiger wäre es, auch bei erwachsenen Patienten eine semiautomatische Methode zur Auswertung des Skelettbefalls mittels PET/CT zur Unterstützung der visuellen Bewertung zu verwenden, wozu die in dieser Arbeit vorgestellte Methode geeignet wäre.

Ebenso könnte der in dieser Arbeit verwendete Algorithmus auf weitere Tumorentitäten adaptiert werden. Das wäre neben Non-Hodgkin-Lymphomen insbesondere für solide Tumoren interessant, bei denen das Ausmaß der skelettalen Metastasierung nachgewiesenermaßen mit der Prognose korreliert. Die bisher vorhandenen Ansätze zur semiautomatischen Auswertung der Tumorlast im PET sind nicht geeignet lediglich skelettale Läsionen zu erfassen, vgl. (Torigian et al. 2011).

Denkbar ist auch eine Anwendung zur Dosimetrie, wenn vor einer Radionuklidtherapie von Skelettmetastasen (z. B.  $^{177}\text{Lu}$ -PSMA-Therapie beim Prostatakarzinom) eine Abschätzung der Strahlendosis des Knochenmarks anhand der vorhandenen Skelettmetastasen erfolgen soll. Eine Eignungsprüfung wird derzeit für ossär metastasierte Patienten mit neuroendokrinen Tumoren und Prostatakarzinomen an der Universitätsklinik Leipzig durchgeführt.

## **8 Zusammenfassung der Arbeit**

### **Hintergrund und Aufgabenstellung**

Für das initiale Staging von Patienten mit Hodgkin-Lymphom (HL) ist die FDG-PET/CT nach internationalem Konsensus die Methode der ersten Wahl (Cheson et al. 2014). Zur Detektion des Skelettbefalls hat die FDG-PET/CT die Knochenmarkbiopsie abgelöst. Der Nachweis eines Skelettbefalls führt zur Einstufung ins Stadium IV nach der Ann-Arbor-Klassifikation und ist somit therapierelevant. Das metabolische Tumervolumen (MTV), das die Gesamtheit aller im Ganzkörper-PET nachweisbaren Tumervolumina darstellt, besitzt prognostische Relevanz. Es wurden bereits verschiedene Algorithmen zur semiautomatischen Bestimmung des MTV entwickelt, diese beziehen sich in der Regel allerdings nur auf die nodalen Befälle.

Im Rahmen dieser Arbeit wurde ein Algorithmus zur semiautomatischen Detektion und Quantifizierung von Mehranreicherungsherden im Skelett zur Erleichterung der Diagnostik des Skelettbefalls, zur standardisierten Auswertung und Erfassung des skelettalen Tumervolumens entwickelt und getestet. Der Einfluss des Ausmaßes des skelettalen Tumervolumens auf das ereignisfreie Überleben und Gesamtüberleben wurde untersucht.

### **Methodik**

Die Untersuchung erfolgte an PET/CT-Datensätzen von Kindern und Jugendlichen mit HL, die im Zeitraum 2007 bis 2013 in die EuroNet-PHL-C1-Studie eingeschlossen worden waren. Bei 142 Kindern wurde im Rahmen der visuellen zentralen Referenzbeurteilung im PET/CT ein Skelettbefall diagnostiziert und dokumentiert. Der weit überwiegende Teil der Patienten wies dagegen keinen Skelettbefall auf und zeigte entweder eine physiologische Aktivitätsbelegung oder aber eine diffuse Mehrbelegung des Skeletts im Sinne einer paraneoplastischen Knochenmarkaktivierung. Für die Testung der verschiedenen methodischen Lösungsansätze wurden jeweils Stichproben aus den Teilkohorten mit Skelettbefall, mit Knochenmarkaktivierung und mit physiologischem skelettalem FDG-Uptake herangezogen. Der beste gefundene Algorithmus wurde schließlich an allen 142 Patienten mit Skelettbefall bezüglich seiner diagnostischen Genauigkeit getestet.

Im ersten semiautomatischen Bearbeitungsschritt wurden die dem Skelett zuzuordnenden Bildpunkte automatisch aus dem CT-Datensatz anhand der Hounsfielddichte (Schwellenwert  $> 160$  HU) segmentiert. Fehlstellen im Knochenmark konnten mit einem Filteralgorithmus gefüllt werden. Diese Skelettmaske wurde auf den PET-Datensatz übertragen.

Es wurden folgende Ansätze zur Detektion eines Skelettbefalls in den PET-Datensätzen getestet:

- 1) Das mittlere Nuklidspeicherverhalten im gesamten, virtuell extrahierten Skelett wurde auf das mittlere Speicherverhalten im Referenzorgan Leber normiert.
- 2) In einer Histogrammanalyse wurden Dichtekurven der Zählraten aller zur Skelettmaske zugehörigen Voxel erstellt. Es wurde überprüft, ob eine Unterscheidung zwischen physiologischem und pathologischem Uptake entweder visuell oder anhand von elf vorimplementierten Histogrammkenngößen gelingt.
- 3) Für eine regionenspezifische Auswertung wurde das Skelett in die sechs Abschnitte Wirbelsäule, Schädel, Rippenthorax, Becken, Humeri und Femora unterteilt. Es wurde geprüft, ob regionenspezifische Schwellenwerte der Aktivitätsbelegung die automatische Detektion von Skelettbefall verbessern.
- 4) Für einen Vergleich der Aktivitätsbelegung von Wirbelkörpern ohne und mit Wirbelkörperläsionen wurden zunächst die vier Voxel mit der höchsten Aktivitätsbelegung in 20 physiologisch anreichernden Wirbelkörpern bestimmt. Anschließend wurden das Volumen von 157 Befallsläsionen in Wirbelkörpern und deren vier Voxel mit der höchsten Aktivitätsbelegung ermittelt und mit den physiologischen Wirbelkörpern verglichen.
- 5) Bei elf zufällig ausgewählten Patienten mit ausgeprägtem Skelettbefall wurde überprüft, ob als Referenzregion die gesamte Skelettmaske oder ein physiologischer Wirbelkörper besser zur Detektion von Skelettbefall geeignet ist. Das Skelett wurde zur Läsionserfassung in 45 einzelne Regionen unterteilt.

Hierbei erwies sich der einzelne Wirbelkörper dem Gesamtskelett als Referenzregion überlegen. Im nächsten Schritt wurde der beste Schwellenwert zur Detektion von Skelettbefall, ausgedrückt in Standardabweichungen (SD) über dem Mittelwert der Zählraten im Referenzwirbel, ausgetestet und eine Mindestgröße der zu detektierenden Läsionen festgelegt, um eine akzeptable Spezifität der Methode zu erzielen.

#### Abschließendes Ergebnis nach Methodentestung – Bester Algorithmus

Nach systematischer Testung unterschiedlicher Parameter wurden Referenzregion (Wirbelkörper) und Schwellenwerte (Aktivitätsbelegung und Mindestgröße) festgelegt. Als Referenzregion wird eine 15 bis 30 ml große VOI in einem nach visueller Einschätzung nicht tumorbefallenen Wirbelkörper (meist LWK 4) platziert. Das Suchprogramm identifiziert skelettale Herde mit einem  $SUV > (SUV_{\text{mean}} + 2,5 \text{ SD})$  der Referenzregion, sofern das Volumen  $> 0,25 \text{ ml}$  ist.

Das nach den eben genannten Kriterien arbeitende Programm wurde an den PET/CT-Untersuchungen der 142 Patienten mit visuell detektiertem Skelettbefall validiert.

#### **Ergebnisse**

Folgende Erkenntnisse konnten aus den einzelnen Ansätzen gewonnen werden:

- 1) Die Aktivitätsbelegung der gesamten Skelettmaske von Patienten mit ausgeprägtem Skelettbefall und physiologischem Skelett lässt sich sicher unterscheiden. Dahingegen lässt sich ein geringer residueller Befall nicht von einem tumorfreien Skelett unterscheiden. Die Aktivitätsbelegung bei Knochenmarkaktivierung zeigt sehr unterschiedliche Ausprägungen, die jedoch nicht sicher von Skelettbefall zu unterscheiden sind.
- 2) Es gelang mittels Histogrammanalyse aufgrund breiter Überlappungsbereiche keine visuelle Differenzierung von Skelettmasken mit Lymphombefall und Knochenmarkaktivierung. Alle elf getesteten Histogrammkenngrößen waren in den durchgeführten Stichproben nicht zur sicheren Unterscheidung zwischen Skeletten mit physiologischem und pathologischem Uptake geeignet.

- 3) Es zeigte sich eine inhomogene Aktivitätsbelegung des physiologischen Skeletts mit systematisch höherer Aktivitätsbelegung von Becken und Wirbelsäule im Vergleich zu Thoraxskelett, Femora und Humeri, für die jeweils verschiedene Schwellenwerte ermittelt werden könnten. Ein einheitlicher Schwellenwert für das gesamte Skelett müsste hingegen so gewählt werden, dass pathologische Mehrbelegungen der Skelettabschnitte mit geringerer Impulsdichte wie Thoraxskelett und Extremitäten sicher erkannt werden, ohne dass die physiologisch höheren Aktivitätsbelegungen der Skelettabschnitte Becken und Wirbelsäule als falsch-positive Läsionen des Hodgkin-Lymphoms fehlgedeutet werden.

Da der SUV des Schädels durch Spillover des Gehirns falsch erhöht gemessen wurde und der Schädel nur in äußerst seltenen Fällen beim Hodgkin-Lymphom befallen ist, wurde der Schädel aus der Skelettmaske entfernt.

- 4) Es zeigte sich, dass einzelne Voxel in nicht befallenen Wirbelkörpern dennoch Aktivitätsbelegungen in der gleichen Höhe wie Lymphommanifestationen annehmen können, somit gibt es einen Überlappungsbereich zwischen physiologischen und pathologischen Aktivitätsbelegungen.
- 5) Die Anwendung eines physiologischen Wirbelkörpers als Referenzregion stellte sich gegenüber der gesamten Skelettmaske insbesondere in Hinblick auf die Sensitivität als überlegen heraus.

Der schlussendlich als am besten geeignete Algorithmus detektierte Skelettbefall bei 130 von 142 Patienten (Sensitivität auf Patientenebene von 91,5 Prozent, wenn die visuelle Beurteilung durch Experten als Goldstandard angenommen wurde). Von 1015 visuell erfassten skelettalen Läsionen wurden 774 durch das Suchprogramm detektiert (Sensitivität auf Regionenebene von 76,3 Prozent). Von den 5375 nicht befallenen Regionen wurde in 5137 Regionen auch keine Läsion vom Algorithmus angezeigt (Spezifität auf Regionenebene von 95,6 Prozent). Bei 91 der 142 Patienten wurde neben den im Dokumentationsbogen der Referenzbegutachtung für Skelettbefall vermerkten Hodgkinmanifestationen keine weitere, tatsächlich intraskelettale Läsion detektiert (Spezifität auf Patientenebene von 64,1 Prozent). Zusätzlich wurden 104 Läsionen gefunden, die in der visuellen Referenzbegutachtung nicht als Skelettbefall bewertet wurden, was 0,73 Läsionen pro Patienten entspricht. Sowohl die vom

Programm nicht erfassten als auch die zusätzlich erfassten Läsionen wurden nach Lage, Größe und Aktivitätsbelegung analysiert. Das Ausmaß des Skelettbefalls wurde mit dem ereignisfreien und Gesamtüberleben verglichen, wobei sich keine Korrelation ergab.

## **Diskussion**

Es wurde ein semiautomatischer Algorithmus entwickelt, der zum Screening auf Skelettbefall in einer Multi-Center-Studie geeignet ist. Er bietet eine objektive, reproduzierbare Vorauswahl von Läsionen, welche bestätigt oder verworfen werden müssen. Durch Auswertung der von der visuellen Bewertung abweichenden Herdbefunde sollen in einem nächsten Schritt der automatische Algorithmus und ggf. auch die visuelle Bewertung optimiert werden. Das individuelle Skelettbefallsvolumen hat sich bei pädiatrischen Hodgkin-Lymphom-Patienten nicht als unabhängiger Risikofaktor für das ereignisfreie oder Gesamtüberleben herausgestellt. Denkbar ist neben einer Anwendung des Algorithmus bei erwachsenen Patienten mit Hodgkin-Lymphom auch ein Einsatz zur Erfassung von Skelettbefall bei Patienten mit Non-Hodgkin-Lymphomen oder soliden Tumoren im Rahmen einer PET/CT-Untersuchung, was jedoch eine Überprüfung und ggf. Modifikation der Schwellenwerte voraussetzt.



## 9 Literaturverzeichnis

Adams, H. J. A.; Kwee, T. C.; Keizer, B. de; Fijnheer, R.; de Klerk, J M H; Littooij, A. S.; Nievelstein, R. A. J. (2014): Systematic review and meta-analysis on the diagnostic performance of FDG-PET/CT in detecting bone marrow involvement in newly diagnosed Hodgkin lymphoma: is bone marrow biopsy still necessary? In: *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 25 (5), S. 921–927. DOI: 10.1093/annonc/mdt533.

Adams, Hugo J. A.; Kwee, Thomas C.; Vermoolen, Malou A.; Keizer, Bart de; de Klerk, John M H; Adam, Judit A. et al. (2013): Whole-body MRI for the detection of bone marrow involvement in lymphoma: prospective study in 116 patients and comparison with FDG-PET. In: *European radiology* 23 (8), S. 2271–2278. DOI: 10.1007/s00330-013-2835-9.

Anand, Aseem; Morris, Michael J.; Kaboteh, Reza; Bath, Lena; Sadik, May; Gjertsson, Peter et al. (2016a): Analytic Validation of the Automated Bone Scan Index as an Imaging Biomarker to Standardize Quantitative Changes in Bone Scans of Patients with Metastatic Prostate Cancer. In: *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* 57 (1), S. 41–45. DOI: 10.2967/jnumed.115.160085.

Anand, Aseem; Morris, Michael J.; Larson, Steven M.; Minarik, David; Josefsson, Andreas; Helgstrand, John T. et al. (2016b): Automated Bone Scan Index as a quantitative imaging biomarker in metastatic castration-resistant prostate cancer patients being treated with enzalutamide. In: *EJNMMI research* 6 (1), S. 23. DOI: 10.1186/s13550-016-0173-z.

Arens, Anne I. J.; Troost, Esther G. C.; Hoeben, Bianca A. W.; Grootjans, Willem; Lee, John A.; Gregoire, Vincent et al. (2014): Semiautomatic methods for segmentation of the proliferative tumour volume on sequential FLT PET/CT images in head and neck carcinomas and their relation to clinical outcome. In: *European journal of nuclear medicine and molecular imaging* 41 (5), S. 915–924. DOI: 10.1007/s00259-013-2651-0.

Bagci, U; Udupa, J K; Yao, J; Mollura, D J (2012): Co-segmentation of functional and anatomical images. In: *Med Image Comput Comput Assist Interv* 15 (3), S. 459–467.

Barrett, J. F.; Keat, N. (2004): Artifacts in CT: Recognition and Avoidance. In: *Radiographics* 24 (6), S. 1679–1691. DOI: 10.1148/rg.246045065.

Beare, Richard (2011): Histogram-based thresholding - some missing methods. In: *The Insight Journal* (6).

Blake, G M; Park-Holohan, S J; Cook, G J; Fogelman, I (2001): Quantitative studies of bone with the use of <sup>18</sup>F-fluoride and <sup>99m</sup>Tc-methylene diphosphonate. In: *Semin Nucl Med.* 31 (1), S. 28–49.

Blebea, Judy S.; Houseni, Mohamed; Torigian, Drew A.; Fan, Chengzhong; Mavi, Ayse; Zhuge, Ying et al. (2007): Structural and functional imaging of normal bone marrow and evaluation of its age-related changes. In: *Seminars in Nuclear Medicine* 37 (3), S. 185–194. DOI: 10.1053/j.semnuclmed.2007.01.002.

Boellaard, Ronald; O'Doherty, Mike J.; Weber, Wolfgang A.; Mottaghy, Felix M; Lonsdale, Markus N.; Stroobants, Sigrid G. et al. (2010): FDG PET and PET/CT:

EANM procedure guidelines for tumour PET imaging: version 1.0. In: *Eur J Nucl Med Mol Imaging* (2007) 34: 60–67 37, S. 181–200.

Boellaard, Ronald; Oyen, Wim J. G.; Hoekstra, Corneline J.; Hoekstra, Otto S.; Visser, Eric P.; Willemsen, Antoon T. et al. (2008): The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. In: *Eur J Nucl Med Mol Imaging* 35 (12), S. 2320–2333.

Brown, Colin; Dempsey, Mary-Frances; Gillen, Gerry; Elliott, Alex T. (2010): Investigation of 18F-FDG 3D mode PET image quality versus acquisition time. In: *Nucl Med Commun* 31 (3), S. 254–259.

Buhmann, Kirchhoff S.; Becker, C.; Duerr, H. R.; Reiser, M.; Baur-Melnyk, A. (2009): Detection of osseous metastases of the spine: comparison of high resolution multi-detector-CT with MRI. In: *European journal of radiology* 69 (3), S. 567–573.

Castell, F.; Cook, G. J. R. (2008): Quantitative techniques in 18FDG PET scanning in oncology. In: *Br J Cancer* 98 (10), S. 1597–1601.

Ceriani, Luca; Martelli, Maurizio; Zinzani, Pier Luigi; Ferreri, Andrés J. M.; Botto, Barbara; Stelitano, Caterina et al. (2015): Utility of baseline 18FDG-PET/CT functional parameters in defining prognosis of primary mediastinal (thymic) large B-cell lymphoma. In: *Blood* 126 (8), S. 950–956. DOI: 10.1182/blood-2014-12-616474.

Cheng, Gang; Chen, Wengen; Chamroonrat, Wichana; Torigian, Drew A.; Zhuang, Hongming; Alavi, Abass (2011): Biopsy versus FDG PET/CT in the initial evaluation of bone marrow involvement in pediatric lymphoma patients. In: *Eur J Nucl Med Mol Imaging* 38 (8), S. 1469–1476.

Cottreau, A. S.; Becker, S.; Broussais, F.; Casasnovas, O.; Kanoun, S.; Roques, M. et al. (2016): Prognostic value of baseline total metabolic tumor volume (TMTV0) measured on FDG-PET/CT in patients with peripheral T-cell lymphoma (PTCL). In: *Annals of oncology : official journal of the European Society for Medical Oncology* 27 (4), S. 719–724. DOI: 10.1093/annonc/mdw011.

Daw, Stephen; Wynn, Rob; Wallace, Hamish (2010): Management of relapsed and refractory classical Hodgkin lymphoma in children and adolescents (152), S. 249–260.

DGN-Geschäftsstelle: DGN-Handlungsempfehlung 031-049, zuletzt geprüft am 31.12.2017.

Dörffel, W.; Riepenhausen, M.; Ludwig, Wolf-Dieter; Schellong, G. (2010): Langzeitfolgen nach Therapie eines Hodgkin-Lymphoms bei Kindern und Jugendlichen. In: *Journal Onkologie* (9), S. 449–456.

El-Galaly, T. C.; d'Amore, F.; Mylam, K. J.; Nully Brown, P. de; Bogsted, M.; Bukh, A. et al. (2012): Routine Bone Marrow Biopsy Has Little or No Therapeutic Consequence for Positron Emission Tomography/Computed Tomography-Staged Treatment-Naive Patients With Hodgkin Lymphoma. In: *Journal of Clinical Oncology* 30 (36), S. 4508–4514.

Elsner, A.; Kurch, L.; Tchavdarova, L.; Georgi, T.; Barthel, A.; Körholz, D. et al. (2012): "Tumor-finder" - A semiautomatic algorithm for detection of tumor lesions in lymphoma. In: *Nuklearmedizin* (2), S. 20.

Elstrom, Rebecca L.; Tsai, Donald E.; Vergilio, Jo-Anne; Downs, Lisa H.; Alavi, Abass; Schuster, Stephen J. (2004): Enhanced marrow 18Ffluorodeoxyglucose

- uptake related to myeloid hyperplasia in Hodgkin's lymphoma can simulate lymphoma involvement in marrow. In: *Clin Lymphoma* 5 (1), S. 62–64.
- Erdi, Yusuf; Humm, John L.; Imbriaco, Massimo; Yeung, Henry; Larson, Steven M (1997): Quantitative Bone Metastases Analysis Based on Image Segmentation. In: *J Nucl Med* 38 (9), S. 1401–1406.
- Evans, WestleyC; Gilmore, David; English, Jeffrey (2011): The Role of PET and PET/CT in Managing the Care of Lymphoma Patients. In: *J Nucl Med Technol* 39 (3), S. 190–194.
- Even-Sapir, Einat (2005): Imaging of malignant bone involvement by morphologic, scintigraphic, and hybrid modalities. In: *J Nucl Med* 46 (8), S. 1356–1367.
- Feltl, D.; Markova, J.; Mocikova, H.; Dedekova, K.; Kozak, T. (2008): Prognostic impact of bone involvement in Hodgkin lymphoma. In: *Neoplasma* 55 (2), S. 96–100.
- Flehsig, Paul; Walker, Christina; Kratochwil, Clemens; König, Laila; Iagura, Andrei; Moltz, Jan et al. (2017): Role of CT Density in PET/CT-Based Assessment of Lymphoma. In: *Molecular imaging and biology : MIB : the official publication of the Academy of Molecular Imaging*. DOI: 10.1007/s11307-017-1155-x.
- Ganeshan, B.; Miles, K. A.; Babikir, S.; Shortman, R.; Afaq, A.; Ardeschna, K. M. et al. (2017): CT-based texture analysis potentially provides prognostic information complementary to interim fdg-pet for patients with hodgkin's and aggressive non-hodgkin's lymphomas. In: *Eur Radiol* 27 (3), S. 1012–1020. DOI: 10.1007/s00330-016-4470-8.
- Gates, GF (1998): SPECT bone scanning of the spine. In: *Seminars in Nuclear Medicine* 28 (1), S. 78–94.
- Georgi, Thomas Walter; Kluge, Regine; Kurch, Lars; Chavdarova, Lidia; Hasenclever, Dirk; Stoevesandt, Dietrich et al. (2018): FDG-PET response of skeletal (bone marrow and bone) involvement after induction chemotherapy in pediatric Hodgkin lymphoma - Are specific response criteria required? In: *J Nucl Med*.
- Hamaoka, T.; Madewell, J. E.; Podoloff, D. A.; Hortobagyi, G. N.; Ueno, N. T. (2004): Bone imaging in metastatic breast cancer. In: *Journal of Clinical Oncology* 22 (14), S. 2942–2953.
- Hasenclever, Dirk; Diehl, Volker; Armitage, James O.; Assouline, David; Björkholm, Magnus; Brusamolino, Ercole et al. (1998): A Prognostic Score for Advanced Hodgkin's Disease. In: *New England Journal of Medicine* 339, S. 1506–1514.
- Hasenclever, Dirk; Kurch, Lars; Mauz-Korholz, Christine; Elsner, Andreas; Georgi, Thomas; Wallace, Hamish et al. (2014): qPET - a quantitative extension of the Deauville scale to assess response in interim FDG-PET scans in lymphoma. In: *European journal of nuclear medicine and molecular imaging* 41 (7), S. 1301–1308. DOI: 10.1007/s00259-014-2715-9.
- Hines-Thomas, Melissa R.; Howard, Scott C.; Hudson, Melissa M.; Krasin, Matthew J.; Kaste, Sue C.; Shulkin, Barry L.; Metzger, Monika L. (2010): Utility of bone marrow biopsy at diagnosis in pediatric Hodgkin's lymphoma. In: *Haematologica* 95 (10), S. 1691–1696.
- Hoekstra, C. J.; Paglianiti, I.; Hoekstra, O. S.; Smit, E. F.; Postmus, P. E.; Teule, G. J. J.; Lammertsma, A. A. (2000): Monitoring response to therapy in cancer using [18F]-2-fluoro-2-deoxy-D-glucose and positron emission tomography: an overview of

different analytical methods. In: *Eur J Nucl Med (European journal of nuclear medicine)* 27 (6), S. 731–743.

Hofer, Matthias (2010): CT-Kursbuch. Ein Arbeitsbuch für den Einstieg in die Computertomographie. 7. Aufl. Düsseldorf: Didamed Verlag GmbH.

Ilica, AhmetTuran; Kocacelebi, Kenan; Savas, Recep; Ayan, Asli (2011): Imaging of Extranodal Lymphoma With PET/CT. In: *Clin Nucl Med* 36 (10), S. 127–138.

Imbriaco, M.; Larson, S. M.; Yeung, H. W.; Mawlawi, O. R.; Erdi, Y.; Venkatraman, E. S.; Scher, H. I. (1998): A new parameter for measuring metastatic bone involvement by prostate cancer: the Bone Scan Index. In: *Clin Cancer Res* 4 (7), S. 1765–1772.

Itti, E.; Juweid, M. E.; Haïoun, C.; Yeddes, I.; Hamza-Maaloul, F.; El Bez, I. et al. (2010): Improvement of Early 18F-FDG PET Interpretation in Diffuse Large B-Cell Lymphoma: Importance of the Reference Background. In: *Journal of Nuclear Medicine* 51 (12), S. 1857–1862. DOI: 10.2967/jnumed.110.080556.

Jung, C. M.; Kugel, H.; Schulte, O.; Heindel, W. (2000): Proton-MR spectroscopy of the spinal bone marrow. An analysis of physiological signal behavior. In: *Radiologe* 40 (8), S. 694–699.

Kanoun, Salim; Rossi, Cédric; Berriolo-Riedinger, Alina; Dygai-Cochet, Inna; Cochet, Alexandre; Humbert, Olivier et al. (2014): Baseline metabolic tumour volume is an independent prognostic factor in Hodgkin lymphoma. In: *Eur J Nucl Med Mol Imaging* 41 (9), S. 1735–1743. DOI: 10.1007/s00259-014-2783-x.

Kero, Tanja; Lindsjo, Lars; Sorensen, Jens; Lubberink, Mark (2015): Accurate analysis and visualization of cardiac 11C-PIB uptake in amyloidosis with semiautomatic software. In: *Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology*. DOI: 10.1007/s12350-015-0149-9.

Keyes, John W. (1995): SUV: Standard Uptake or Silly Useless Value? In: *J Nucl Med (Journal of nuclear medicine)* 36 (10), S. 1836–1839.

Kluge, R.; Körholz, D. (2011): Die Bedeutung des FDG-PET für die Stadieneinteilung und Therapie des Hodgkin-Lymphoms im Kindesalter. In: *Klinische Pädiatrie* (223), S. 315–319.

Kurch, L.; Mauz-Korholz, C.; Bertling, S.; Wallinder, M.; Kaminska, M.; Marwede, D. et al. (2013a): The EuroNet paediatric hodgkin network - modern imaging data management for real time central review in multicentre trials. In: *Klinische Padiatrie* 225 (6), S. 357–361. DOI: 10.1055/s-0033-1354416.

Kurch, L.; Sabri, O.; Kluge, R. (2013b): Die Rolle der FDG-PET/CT bei malignen Lymphomen im Kindes- und Jugendalter unter besonderer Berücksichtigung des Hodgkin-Lymphoms. In: *Der Nuklearmediziner* (36), S. 96–106.

Kurch, Lars (2011): Analyse des semiquantitativen FDG-PET-Parameters "SUVmax" im Rahmen einer multizentrischen Pilotstudie zu Hodgkinlymphomen im Kindes- und Jugendalter. Universität Leipzig, Leipzig. Klinik und Poliklinik für Nuklearmedizin.

Larson, Steven M. (2016): EXINI Quantitative Bone Scan Index: Expanded Utility for the Planar Radionuclide Bone Scan. In: *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* 57 (1), S. 5–6. DOI: 10.2967/jnumed.115.164137.

Liang, Mingzhu; Tang, Wei; Xu, Dong Ming; Jirapatnakul, Artit C.; Reeves, Anthony P.; Henschke, Claudia I.; Yankelevitz, David (2016): Low-Dose CT Screening for

Lung Cancer: Computer-aided Detection of Missed Lung Cancers. In: *Radiology*, S. 150063. DOI: 10.1148/radiol.2016150063.

Lopci, Egesta; Burnelli, Roberta; Guerra, Luca; Cistaro, Angelina; Piccardo, Arnoldo; Zucchetto, Pietro et al. (2011): Postchemotherapy PET evaluation correlates with patient outcome in paediatric Hodgkin's disease. In: *Eur J Nucl Med Mol Imaging* 38 (9), S. 1620–1627.

Martin, Peter (2009): Is there a role for bone scan in Hodgkin lymphoma in the age of 18 Is there a role for bone scan in Hodgkin lymphoma in t 18F-fluorodeoxyglucose positron emission tomography? In: *Leukemia & Lymphoma* 50 (11), S. 1731–1732.

Mauz-Korholz, C.; Hasenclever, D.; Dorffel, W.; Ruschke, K.; Pelz, T.; Voigt, A. et al. (2010): Procarbazine-Free OEPA-COPDAC Chemotherapy in Boys and Standard OPPA-COPP in Girls Have Comparable Effectiveness in Pediatric Hodgkin's Lymphoma: The GPOH-HD-2002 Study. In: *Journal of Clinical Oncology* 28 (23), S. 3680–3686.

Meignan, Michel; Gallamini, Andrea; Haioun, Corinne (2009): Report on the First International Workshop on Interim-PET-Scan in Lymphoma. In: *Leukemia & Lymphoma* 50 (8), S. 1257–1260. DOI: 10.1080/10428190903040048.

Meignan, Michel; Itti, Emmanuel; Gallamini, Andrea; Younes, Anas (2015): FDG PET/CT imaging as a biomarker in lymphoma. In: *Eur J Nucl Med Mol Imaging* 42 (4), S. 623–633. DOI: 10.1007/s00259-014-2973-6.

Metwally, Hussein; Courbon, Frédéric; David, Isabelle; Filleron, Thomas; Blouret, Aurélien; Rives, Michel et al. (2011): Coregistration of Prechemotherapy PET-CT for Planning Pediatric Hodgkin's Disease Radiotherapy Significantly Diminishes Interobserver Variability of Clinical Target Volume Definition. In: *International Journal of Radiation Oncology\*Biophysics* 80 (3), S. 793–799.

Miles, Kenneth A.; Ganeshan, Balaji; Hayball, Michael P. (2013): CT texture analysis using the filtration-histogram method. What do the measurements mean? In: *Cancer imaging : the official publication of the International Cancer Imaging Society* 13 (3), S. 400–406. DOI: 10.1102/1470-7330.2013.9045.

Murata, Yuji; Kubota, Kazuo; Yukihiro, Masashi; Ito, Kimiteru; Watanabe, Hiroshige; Shibuya, Hitoshi (2006): Correlations between 18F-FDG uptake by bone marrow and hematological parameters: measurements by PET/CT. In: *Nuclear Medicine and Biology* 33 (8), S. 999–1004.

Muzahir, S.; Mian, M.; Munir, I.; K Nawaz, M.; S Faruqui, Z.; A Mufti, K. et al. (2012): Clinical utility of 18F FDG-PET/CT in the detection of bone marrow disease in Hodgkin's lymphoma. In: *British Journal of Radiology* 85 (1016), S. 490–496.

Nakajo, Masatoyo; Nakajo, Masayuki; Jinguji, Megumi; Fukukura, Yoshihiko; Nakabeppu, Yoshiaki; Tani, Atsushi; Yoshiura, Takashi (2015): The value of intratumoral heterogeneity of (18)F-FDG uptake to differentiate between primary benign and malignant musculoskeletal tumours on PET/CT. In: *The British journal of radiology* 88 (1055), S. 20150552. DOI: 10.1259/bjr.20150552.

Nehmeh, Sadek A.; Erdi, Yusuf E. (2008): Respiratory motion in positron emission tomography/computed tomography: a review. In: *Seminars in Nuclear Medicine* 38 (3), S. 167–176. DOI: 10.1053/j.semnuclmed.2008.01.002.

Olson, Michael R.; Donaldson, Sarah S. (2008): Treatment of pediatric hodgkin lymphoma. In: *Curr Treat Options Oncol* 9 (1), S. 81–94.

- Paulino, Arnold C.; Margolin, Judith; Dreyer, ZoAnn; Teh, Bin S.; Chiang, Stephen (2012): Impact of PET-CT on involved field radiotherapy design for pediatric Hodgkin lymphoma. In: *Pediatr. Blood Cancer* 58 (6), S. 860–864. DOI: 10.1002/pbc.23273.
- Pelosi, E.; Penna, D.; Deandreis, D.; Chiappella, A.; Skanjeti, A.; Vitolo, U.; Bisi, G. (2008): Bone marrow disease detection with FDG-PET/CT and bone marrow biopsy during the staging of malignant lymphoma: results from a large multicentre study. In: *Q J Nucl Med Mol Imaging* 52 (1), S. 9–16.
- Poeppel, T. D.; Krause, B. J.; Heusner, T. A.; Boy, C.; Bockisch, A.; Antoch, G. (2009): PET/CT for the staging and follow-up of patients with malignancies. In: *Eur J Radiol* 70 (3), S. 382–392.
- Puri, Tanuj; Blake, Glen M.; Curran, Kathleen M.; Carr, Hamish; Moore, Amelia E. B.; Colgan, Niall et al. (2012): Semiautomatic region-of-interest validation at the femur in (18)F-fluoride PET/CT. In: *Journal of nuclear medicine technology* 40 (3), S. 168–174. DOI: 10.2967/jnmt.111.100107.
- Purz, Sandra; Mauz-Körholz, Christine; Körholz, Dieter; Hasenclever, Dirk; Krause, Antje; Sorge, Ina et al. (2011): [18F]Fluorodeoxyglucose Positron Emission Tomography for Detection of Bone Marrow Involvement in Children and Adolescents With Hodgkin's Lymphoma. In: *J Clin Oncol* 29 (26).
- Rathore, N.; Eissa, H. M.; Margolin, J. F.; Liu, H.; Wu, M. F.; Horton, T. et al. (2012): Pediatric Hodgkin lymphoma: are we over-scanning our patients? In: *Pediatr Hematol Oncol* 29 (5), S. 415–423.
- Robertson, Virginia Lang; Anderson, Cynthia S.; Keller, Frank G.; Halkar, Raghuveer; Goodman, Michael; Marcus, Robert B.; Esiashvili, Natia (2011): Role of FDG-PET in the definition of involved-field radiation therapy and management for pediatric Hodgkin's lymphoma. In: *Int J Radiat Oncol Biol Phys* 80 (2), S. 324–332.
- Rodríguez-Vigil, Beatriz; Gómez-León, Nieves; Pinilla, Inmaculada; Hernández-Maraver, Dolores; Coya, Juan; Martín-Curto, Luis (2006): Positron Emission Tomography/Computed Tomography in the Management of Hodgkin's Disease and Non-Hodgkin's Lymphoma. In: *Current Problems in Diagnostic Radiology* 35 (4), S. 151–163.
- Rosen, Ron S.; Fayad, Laura; Wahl, Richard L. (2006): Increased 18F-FDG uptake in degenerative disease of the spine: Characterization with 18F-FDG PET/CT. In: *J Nucl Med* 47 (8), S. 1274–1280.
- Ruhlmann, J; Oeler, P; Breisack, HJ (Hg.) (1998): PET in der Onkologie. 1. Aufl. Heidelberg: Springer Verlag.
- Sambuceti, Gianmario; Brignone, Massimo; Marini, Cecilia; Massollo, Michela; Fiz, Francesco; Morbelli, Silvia et al. (2012): Estimating the whole bone-marrow asset in humans by a computational approach to integrated PET/CT imaging. In: *Eur J Nucl Med Mol Imaging* 39 (8), S. 1326–1338.
- Shreve, P. D.; Anzai, Y.; Wahl, R. L. (1999): Pitfalls in oncologic diagnosis with FDG PET imaging. Physiologic and benign variants. In: *Radiographics* 19 (1), 61-77; quiz 150-1. DOI: 10.1148/radiographics.19.1.g99ja0761.
- Shulkin, Barry L.; Goodin, Geoffrey S.; McCarville, Mary B.; Kaste, Sue C.; Krasin, Matthew J.; Hudson, Melissa M.; Metzger, Monika (2009): Bone and [18F]Fluorodeoxyglucose positron-emission tomography/computed tomography

scanning for the assessment of osseous involvement in Hodgkin lymphoma in children and young adults. In: *Leukemia & Lymphoma* 50 (50), S. 1794–1802.

Simpson, C. David; Gao, Jun; Fernandez, Conrad V.; Yhap, Margaret; Price, Victoria E.; Berman, Jason N. (2008): Routine bone marrow examination in the initial evaluation of paediatric Hodgkin lymphoma: the Canadian perspective. In: *Br J Haematol* 141 (6), S. 820–826.

Sjostrand, Karl; Ohlsson, Mattias; Edenbrandt, Lars (2009): Statistical regularization of deformation fields for atlas-based segmentation of bone scintigraphy images. In: *Med Image Comput Comput Assist Interv* 12 (1), S. 664–671.

Smith, M. A.; Seibel, N. L.; Altekruse, S. F.; Ries, L. A. G.; Melbert, D. L.; O'Leary, M. et al. (2010): Outcomes for Children and Adolescents With Cancer: Challenges for the Twenty-First Century. In: *Journal of Clinical Oncology* 28 (15), S. 2625–2634.

Soloway, M. S.; Hardeman, S. W.; Hickey, D.; Raymond, J.; Todd, B.; Soloway, S.; Moinuddin, M. (1988): Stratification of patients with metastatic prostate cancer based on extent of disease on initial bone scan. In: *Cancer* 61 (1), S. 195–202.

Song, Moo-Kon; Chung, Joo-Seop; Shin, Ho-Jin; Moon, Joon-Ho; Lee, Jeong-Ok; Lee, Ho-Sup et al. (2012): Prognostic value of metabolic tumor volume on PET / CT in primary gastrointestinal diffuse large B cell lymphoma. In: *Cancer science* 103 (3), S. 477–482. DOI: 10.1111/j.1349-7006.2011.02164.x.

Soret, Marine; Bacharach, Stephen L.; Buvat, Irene (2007): Partial-volume effect in PET tumor imaging. In: *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* 48 (6), S. 932–945. DOI: 10.2967/jnumed.106.035774.

Stauss, J.; Franzius, C.; Pfluger, T.; Juergens, K. U.; Biassoni, L.; Begent, J. et al. (2008): Guidelines for 18F-FDG PET and PET-CT imaging in paediatric oncology. In: *Eur J Nucl Med Mol Imaging* (35), S. 1581–1588.

Tallen, G.; Henze, G; Creutzig, U.; Dworzak, M.; Klingebiel, T. (2010): Auswirkungen der EU-Direktive für klinische Studien auf Kinder und Jugendliche mit Krebserkrankungen in Europa. In: *FORUM* 25, S. 42–48.

Thie, Joseph A. (2004): Understanding the Standardized Uptake Value, Its Methods, and Implications for Usage. In: *THE JOURNAL OF NUCLEAR MEDICINE* 45 (9).

Torigian, Drew A.; Lopez, Rosa Fernandez; Alapati, Sridevi; Bodapati, Geetha; Hofheinz, Frank; van den Hoff, Joerg et al. (2011): Feasibility and performance of novel software to quantify metabolically active volumes and 3D partial volume corrected SUV and metabolic volumetric products of spinal bone marrow metastases on 18F-FDG-PET/CT. In: *Hell J Nucl Med* 14 (1), S. 8–14.

Ulmert, David; Kaboteh, Reza; Fox, Josef J.; Savage, Caroline; Evans, Michael J.; Lilja, Hans et al. (2012): A novel automated platform for quantifying the extent of skeletal tumour involvement in prostate cancer patients using the Bone Scan Index. In: *European urology* 62 (1), S. 78–84. DOI: 10.1016/j.eururo.2012.01.037.

Wahl, Richard L.; Heather, Jacene; Kasamon, Yvette; Lodge, Martin A. (2009): From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors. In: *J Nucl Med (Journal of nuclear medicine)* 50 (5), S. 122–150.

Warburg, Otto (1956): On the origin of cancer cells. In: *Science* 123 (3191).

Watabe, Tadashi; Tatsumi, Mitsuaki; Watabe, Hiroshi; Isohashi, Kayako; Kato, Hiroki; Yanagawa, Masahiro et al. (2012): Intratumoral heterogeneity of F-18 FDG uptake

differentiates between gastrointestinal stromal tumors and abdominal malignant lymphomas on PET/CT. In: *Annals of nuclear medicine* 26 (3), S. 222–227. DOI: 10.1007/s12149-011-0562-3.

Weiler-Sagie, Michal; Kagna, Olga; Dann, Eldad J.; Ben-Barak, Ayelet; Israel, Ora (2014): Characterizing bone marrow involvement in Hodgkin's lymphoma by FDG-PET/CT. In: *Eur J Nucl Med Mol Imaging* 41 (6), S. 1133–1140. DOI: 10.1007/s00259-014-2706-x.

Zasadny, K. R.; Wahl, R. L. (1993): Standardized uptake values of normal tissues at PET with 2-fluorine-18-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. In: *Radiology* 189 (3), S. 847–850.



## 10 Abbildungsverzeichnis

Abb. 1: Deauville-5-Punkte-Skala für PET-Scans (Meignan et al. 2009)	6
Abb. 2: Entwicklung von malignen Zweiterkrankungen nach Jahren der Therapie (Quelle: (Dörffel et al. 2010), SMN=sekundäre maligne Neoplasien, SST=sekundäre solide Tumoren)	9
Abb. 3: Patient mit Knochenmarkaktivierung	14
Abb. 4: Patient mit Läsionen in Wirbelsäule ohne Knochenmarkaktivierung	14
Abb. 5: Patient mit starker Aktivitätsbelegung der Epiphysenfugen der distalen Radii und proximalen Femora	14
Abb. 6: Knochenbefallsbogen	18
Abb. 7: Eindeutige Läsionen in LWK 1 und BWK 12, fragliche Läsion in LWK 4	20
Abb. 8: Benutzeroberfläche des Hodgkin Super Servers mit Software Andreas Hybrid	24
Abb. 9: CT-Skelettmaske vor Einsatz der Knochenfüllerfunktion mit Fehlstellen in Wirbelkörpern	26
Abb. 10: Knochenfüllerfunktion schließt Fehlstellen in CT-Skelettmaske	26
Abb. 11: CT-Skelettmaske wird auf PET übertragen	27
Abb. 12: Kontrastmittel im Darm wird mittels „Crop VOI“-Funktion markiert	27
Abb. 13: CT-Skelettmaske bei 160 HU	28
Abb. 14: CT-Skelettmaske bei 200 HU	28
Abb. 15: CT-Skelettmaske bei 300 HU	28
Abb. 16: Positionierung des 30 ml-Leberquaders	31
Abb. 17: Beispielhistogramm mit Kenngrößen	33
Abb. 18: Histogramm von Pat-ID 3202 ohne Skelettbefall	35
Abb. 19: Histogramm von Pat-ID 3161 ohne Skelettbefall	35
Abb. 20: Histogramm von Pat-ID 3202 ohne Skelettbefall	36
Abb. 21: Histogramm von Pat-ID 3161 ohne Skelettbefall	36
Abb. 22: Darstellung der sechs verschiedenen Skelettregionen zur Aktivitätsanalyse in unterschiedlichen Schnittebenen	38
Abb. 23: Vermessung einzelner Läsionen und Ermittlung $SUV_{peak}$	42
Abb. 24: Markierung des 4. Lendenwirbels mit „Crop VOI“-Funktion	45
Abb. 25: Körper des vierten Lendenwirbels wird in der Transversalrekonstruktion markiert	45
Abb. 26: LWK 4 als Referenzregion in der Sagittalrekonstruktion	45
Abb. 27: Sagittalrekonstruktion einer standardisierten, auf den PET-Datensatz übertragenen Skelettmaske	50
Abb. 28: Koronalrekonstruktion einer standardisierten, auf den PET-Datensatz übertragenen Skelettmaske	50
Abb. 29: Skelett mit Fehlstellen in proximalen Femora und Humeri sowie im rechten Os ilium	55
Abb. 30: Skelett mit Fehlstellen in Os ischium rechts und proximalen Femora	55
Abb. 31: Fehlstelle im rechten oberen Schambeinast durch große Osteolyse	55
Abb. 32: Boxplots der durchschnittlichen skelettalen Aktivitätsbelegung bei unterschiedlichem Knochenmarkstatus	57
Abb. 33: Vergleich $qPET_{mean}$ -Werte bei Patienten mit unterschiedlichem Knochenmarkstatus	57
Abb. 34: Normierter skelettaler Uptake der sieben Patienten im Initialstaging und ERA	60
Abb. 35: Beispielhistogramm 1 ohne Skelettbefall oder Knochenmarkaktivierung	62

Abb. 36: Beispielhistogramm 2 ohne Skelettbefall oder Knochenmarkaktivierung	62
Abb. 37: Beispielhistogramm 1 mit ausgeprägtem Skelettbefall	62
Abb. 38: Beispielhistogramm 2 mit ausgeprägtem Skelettbefall	62
Abb. 39: Beispielhistogramm 1 mit Knochenmarkaktivierung	62
Abb. 40: Beispielhistogramm 2 mit Knochenmarkaktivierung	62
Abb. 41: Boxplots der $qPET_{mean}$ -Werte bei Skelettbefall im Initialstaging, getrennt nach Skelettregionen	66
Abb. 42: Boxplots der $qPET_{mean}$ -Werte im ERA-PET, getrennt nach Skelettregionen	67
Abb. 43: Boxplots der $qPET_{mean}$ -Werte ohne Skelettbefall im Initialstaging, getrennt nach Skelettregionen	67
Abb. 44: Spillover der Hirnaktivität auf die Schädelkalotte	71
Abb. 45: Streudiagramm des normierten $SUV_{peak}$ gegenüber dem jeweiligen Läsionsvolumen	76
Abb. 46: Verteilung der Läsionen auf verschiedene Skelettregionen	77
Abb. 47: Fokale Läsionen in LWK 1 und BWK 12, flau-diffuse Läsion in LWK 4	78
Abb. 48: Patient mit Mehranreicherungen in Epiphysenfugen von Humeri und Femora	81
Abb. 49: Darstellung der Läsionen in der CT-Skelettmaske mit Referenzregion LWK 3 (Sagittalrekonstruktion)	86
Abb. 50: Darstellung der Läsionen in der CT-Skelettmaske mit Referenzregion LWK 3 (Koronalrekonstruktion)	86
Abb. 51: Blasenaktivität projiziert sich über die Schambeinäste (Transversalrekonstruktion)	89
Abb. 52: Blasenaktivität projiziert sich über die Schambeinäste (Koronalrekonstruktion)	89
Abb. 53: Braunes Fettgewebe paravertebral in einer Koronalrekonstruktion dargestellt	90
Abb. 54: Braunes Fettgewebe paravertebral in einer Sagittalrekonstruktion dargestellt	90
Abb. 55: Histogramm des Gesamtvolumens richtig-positiver Läsionen pro Patient	95
Abb. 56: Histogramm des Volumenanteils der richtig-positiven Läsionen an der CT-Skelettmaske	96
Abb. 57: Streudiagramm zwischen Leukozytenzahl und $SUV_{mean}$ des Wirbelkörpers	97
Abb. 58: Streudiagramm der logarithmierten Volumina der Lymphknotenherde und Skelettläsionen	98
Abb. 59: Ereignisfreies Überleben der Patienten abhängig vom quantifizierten Skelettbefall	99

# **11 Tabellenverzeichnis**

Tab. 1: Dichteschwelle (in HU) zur Erstellung der CT-Skelettmaske	54
Tab. 2: Kennwerte der Aktivitätsbelegungen der CT-Skelettmaske bei Patienten mit unterschiedlichem Knochenmarkstatus	56
Tab. 3: qPET-Werte der CT-Skelettmasken bei Patienten mit physiologischem Skelett	58
Tab. 4: qPET-Werte der CT-Skelettmaske bei Patienten mit unterschiedlichem Knochenmarkstatus	59
Tab. 5: SUV- und qPET-Werte der CT-Skelettmaske bei 7 Patienten mit Skelettbefall im Initialstaging und ERA-PET	60
Tab. 6: qPET-Werte der CT-Skelettmaske bei Patienten im ERA-PET ohne und mit geringem residuellem Befall	61
Tab. 7: Position der Analytics mit Kilocounts als Abszisse im Histogramm bei Patienten mit physiologischem Skelett	64
Tab. 8: Position der Analytics mit qPET-Wert als Abszisse im Histogramm bei Patienten mit physiologischem Skelett	64
Tab. 9: Durchschnitt der qPET-Werte, getrennt nach Skelettregionen bei Patienten mit Skelettbefall im Initialstaging	66
Tab. 10: qPET-Werte, getrennt nach Skelettregionen bei Patienten im ERA und Veränderung gegenüber dem Initialstaging	68
Tab. 11: Median und Range der $qPET_{mean}$ -Werte, getrennt nach Skelettregionen bei Patienten mit residuellem Befall im ERA und Veränderung gegenüber dem Initialstaging	69
Tab. 12: Median und Range der $qPET_{max}$ -Werte, getrennt nach Skelettregionen bei Patienten mit residuellem Befall im ERA und Veränderung gegenüber der Initialstaging	69
Tab. 13: Mittlere und maximale SUV-Werte, getrennt nach Skelettregionen bei Patienten ohne Skelettbefall	71
Tab. 14: Mittlere und maximale qPET-Werte, getrennt nach Skelettregionen bei Patienten ohne Skelettbefall	72
Tab. 15: Vergleich 75%-Perzentile $qPET_{mean}$ physiologisches Skelett vs. 25%-Perzentile $qPET_{mean}$ mit Skelettbefall, getrennt nach Skelettregionen	73
Tab. 16: Median und Range der Aktivitätsbelegung der Skelettregionen, getrennt nach Nuklidbelegung des Leberquaders	73
Tab. 17: Median und Range der qPET-Werte der Skelettregionen, getrennt nach Nuklidbelegung des Leberquaders	74
Tab. 18: Kennwerte bei 20 physiologischen Wirbelkörpern	75
Tab. 19: Sensitivität und Spezifität bei Test verschiedener Standardabweichungen über der Aktivitätsbelegung der gesamten CT-Skelettmaske	78
Tab. 20: Sensitivität und Spezifität bei Test verschiedener Standardabweichungen über der Aktivitätsbelegung eines physiologischen Wirbelkörpers	79
Tab. 21: Verhältnis der falsch-positiven Läsionen gegenüber der geprüften Regionenzahl und Volumen falsch-positiver Läsionen, getrennt nach Skelettregionen	80
Tab. 22: Überschreitung der Aktivitätsbelegung der falsch-positiven Läsionen gegenüber dem Schwellenwert des Algorithmus	82
Tab. 23: Vergleich der Sensitivität des Algorithmus bei Anwendung des Größenparameters 4 Voxel vs. 0,25 ml	83
Tab. 24: Referenzwirbelkörper (ohne Skelettbefall) für Suchalgorithmus	85

Tab. 25: Verteilung der richtig-positiven Läsionen auf verschiedene Skelettregionen	87
Tab. 26: Lokalisation und Anteil der Skelettregionen an den zusätzlich detektierten Läsionen	91
Tab. 27: Lokalisation und Anteil der zusätzlich detektierten Läsionen in verschiedenen Skelettregionen	91
Tab. 28: Häufigkeit der Gründe für die Nichtdetektion der falsch-negativen Läsionen	92
Tab. 29: Lokalisation und Anteil der falsch-negativen Läsionen in verschiedenen Skelettregionen	93
Tab. 30: Anzahl der Läsionen bei Patienten mit Skelettbefall ohne durch den Algorithmus detektierte Herde	94
Tab. 31: Stratifikation der Patienten mit Skelettmetastasen bei Prostatakarzinom nach (Soloway et al. 1988)	124

## **12 Erklärung über die eigenständige Abfassung der Arbeit**

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne unzulässige Hilfe oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Ich versichere, dass Dritte von mir weder unmittelbar noch mittelbar eine Vergütung oder geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen, und dass die vorgelegte Arbeit weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde. Alles aus anderen Quellen und von anderen Personen übernommene Material, das in der Arbeit verwendet wurde oder auf das direkt Bezug genommen wird, wurde als solches kenntlich gemacht. Insbesondere wurden alle Personen genannt, die direkt an der Entstehung der vorliegenden Arbeit beteiligt waren. Die aktuellen gesetzlichen Vorgaben in Bezug auf die Zulassung der klinischen Studien, die Bestimmungen des Tierschutzgesetzes, die Bestimmungen des Gentechnikgesetzes und die allgemeinen Datenschutzbestimmungen wurden eingehalten. Ich versichere, dass ich die Regelungen der Satzung der Universität Leipzig zur Sicherung guter wissenschaftlicher Praxis kenne und eingehalten habe.

.....  
Datum

.....  
Unterschrift

## **13 Lebenslauf**

### **Persönliche Daten**

Name	Maximilian Epstude
Geburtsdatum	27. Januar 1988
Geburtsort	Jena
Staatsangehörigkeit	deutsch

### **Beruf**

seit 12/2017	Assistenzarzt für Radiologie am Klinikum Wilhelmshaven
01/2014 – 9/2017	Assistenzarzt für Radiologie in der Ev.-Luth. Diakonissenanstalt zu Flensburg
10/2013 – 12/2013	Introlæge (entspricht einem Assistenzarzt im 1. Jahr) in Sønderburg (Dänemark)

### **Promotionssemester**

10/2012 – 03/2013	Promotionsstipendium der Universität Leipzig
-------------------	--

### **Auslandsaufenthalte**

04/2012 – 07/2012	3. Tertial des Praktischen Jahres am Kantonsspital Luzern (Schweiz), Abteilung für Radiologie und Nuklearmedizin
12/2011 – 04/2012	2. Tertial des Praktischen Jahres an der University of Alabama in Birmingham (USA), Internal Medicine

### **Studium**

10/2008 – 06/2013	Medizinstudium an der Universität Leipzig, klinischer Abschnitt (Abschluss: 2. Staatsexamen, Note: gut)
10/2006 – 09/2008	Medizinstudium an der Charité-Universitätsmedizin Berlin,

vorklinischer Abschnitt (Abschluss: 1.  
Staatsexamen, Note: sehr gut)

## **Schule**

2006

Abitur am Carl-Zeiss-Gymnasium Jena  
mit mathematisch-  
naturwissenschaftlichen Spezialklassen  
(Note 1,0) sowie Spezialschulabitur in  
Mathematik und Chemie

.....  
Datum

.....  
Unterschrift

## **14 Danksagung**

An erster Stelle gilt mein Dank meiner Doktormutter, Frau Prof. Regine Kluge, und meinem Betreuer, Herrn Dr. Lars Kurch, für die stetige wissenschaftliche und methodische Unterstützung während der gesamten Bearbeitungsphase meiner Dissertation.

Frau Prof. Regine Kluge danke ich insbesondere für die Überlassung des Themas und die unkomplizierte Bereitstellung von Daten und Programmen. Herrn Dr. Lars Kurch möchte ich für die vielen fachlichen Gespräche, Ratschläge und Anmerkungen zu meinen Textentwürfen danken. Dieser stetige Austausch hat mich immer wieder neue Aspekte und Ansätze entdecken lassen und mir bei der Strukturierung der Arbeit geholfen. Auch die vielen nicht-wissenschaftlichen und motivierenden Gespräche haben letztendlich zum Gelingen der Arbeit beigetragen.

Mein Dank gilt ebenso der Arbeitsgruppe zur PET-Bildgebung beim pädiatrischen Hodgkin-Lymphom, die mir stets hilfreich und kollegial zur Seite stand. Herrn Thomas Georgi danke ich für die Präsentation der Ergebnisse dieser Arbeit auf der Jahrestagung der Deutschen Gesellschaft für Nuklearmedizin 2013, Herrn Dr. Dirk Hasenclever für die Unterstützung bei der statistischen Auswertung und Frau Ute Rosner für die intensive Lektoratsarbeit.

Außerdem danke ich Herrn Andreas Elsner der Firma Hermes Medical Solutions (Schweden), der durch Modifikationen des vorbestehenden Programmes die Überprüfung verschiedener Ansätze und Implementation des Algorithmus zur Detektion des Skelettbefalls ermöglicht hat.

Ferner möchte ich mich bei der EuroNet-Studiengruppe und ihren Studienleitern, Herrn Prof. Körholz und Frau Prof. Mauz-Körholz von der Universitätsklinik Gießen-Marburg bedanken. Ohne sie gäbe es weder die EuroNet-Studien noch die vorliegende Dissertationsschrift.

Der medizinischen Fakultät der Universität Leipzig danke ich für die Gewährung eines Promotionsstipendiums, was mir den Freiraum verschaffte, mich 6 Monate intensiv der Bearbeitung des Dissertationsthemas zu widmen.



Besonders möchte ich an dieser Stelle meiner Familie danken, vor allem meiner Frau Saskia Epstude für die stete Motivation und Lektoratsarbeit, meinen Eltern Dr. Jörg und Dr. Martina Epstude für ihr großes Interesse am erfolgreichen Gelingen dieser Arbeit sowie meiner Schwester Katharina Epstude und meiner Schwägerin Felicia Möller für ihre Unterstützung bei Literaturrecherchen und statistischen Auswertungen.